LONDON SCHOOL OF ECONOMICS AND POLITICAL SCIENCE

STUDIES IN STATISTICS AND SCIENTIFIC METHOD

Edited by PROFESSOR A. L. BOWLEY AND PROFESSOR A. WOLF

No. 2. Family Expenditure

FAMILY EXPENDITURE

A Study of its Variation

BY

R. G. D. ALLEN, M.A.

AND

A. L. BOWLEY, Sc.D.

STAPLES PRESS LIMITED

P. S. King and Staples Limited

CAVENDISH PLACE, LONDON, W1

PRINTED IN GREAT BRITAIN

PREFACE

THIS investigation is developed from three converging sources. One is an article in *Econometrica*, 1935, "The Action of Economic Forces in Producing Frequency Distributions, etc.," followed by an unpublished communication to the Econometric Society in 1934 on "The Variation of Expenditure." Another is "A Reconsideration of the Theory of Value" in *Economica*, 1934. The third is a group of budgets obtained in connection with the Enquiry into Family Life organised by Sir William Beveridge in co-operation with the B.B.C. When these were analysed, it became evident that their value would be greatly increased if they were brought into relation with other collections, and with the help of Dr. Staehle's list of Budgets, *Econometrica*, 1935, we were able to choose twenty-one groups for comparison. We are indebted to Mr. P. K. O'Brien and Mr. D. Glass for help in the tabulation of the original budgets.

It may be of interest to record that the mathematical formula given at the end of Section 4 of the Appendix, which is in fact fundamental in the analysis, was obtained independently by the two authors in the Autumn of 1933.

THE LONDON SCHOOL OF ECONOMICS.

August, 1935.

CONTENTS

INTRODUCTION

There is no accounting for tastes.
Different people have different opinions;
Some like apples and some like onions.

THE purposes of this study are to discover how far the expenditure of individual families, or of groups of families, can be described by rules and formulæ, to relate any rules that are found to the postulates of economic theory and to describe the variations from the averages that result from the different choices of individual families.

If in the end we cannot account for tastes, we may still be able to account for, and measure, the norms from which they vary. We may find that, though the individual is beyond law, yet the variation shown within a group conforms to a simple rule of distribution, especially if we treat apples and onions as alternative means of satisfying the same desires. Where the Cosmos ends and Chaos begins, we can perhaps extend the scope of science so as to describe even Chaos.

The analysis of demand in economic theory is based on the complex of desires of an individual, and the demand of a group of individuals is derived by a process of summing or averaging the demands of the separate individuals. The question whether the individuals have similar desires is not usually considered, nor indeed is it necessarily relevant to the analysis. It is, however, a question of considerable importance; a change of prices may affect individuals in different ways even if their incomes and responsibilities are the same. It may be possible to discover, from records of expenditure, whether individuals do in fact behave in similar ways when prices or incomes change, or to what extent their behaviour

varies, and so to throw light on the homogeneity of the group in respect of the complex of desires.

The material for such an investigation is to be found in collections of budgets of family expenditure. Hitherto, the use of budget collections has been principally to form bases for the construction of index-numbers showing the changes in the cost of living. A secondary use has been to examine the influence of the level of income on the proportions devoted to different items of the budget. In both cases, attention is paid to the average of defined classes but individual variations within the classes are ignored.

But, if we proceed on the basis of the unaveraged budgets, we can give a more general, and at the same time a more exact, expression to the changing proportions of expenditure. We can also determine whether, as we pass up the scale of incomes, we are still dealing with the same intrinsic tastes, for the satisfaction of which increasing scope comes with increasing income, or whether we pass from one stratum of desires to another.

The postulates of one branch at least of economic theory lead, as we shall show, to a formula for expenditure on any defined item of the budget. When the only variable is income and when the complex of desires is the same for all members of the class considered, this formula takes, as a first approximation, the form

$$y = ke + c$$

where e measures the total expended income, y the expenditure on a particular item and k and c are constants. If the determination of k and c differs from one group to another, then these groups are not alike in respect of desires for the item in question.

The present study is based on the following plan:

(1) Existing budget collections in the usual form of averages are examined to determine whether the expenditures of the group concerned satisfy this or some other simple formula. We can, in this way, throw light on

their validity as bases for the construction of cost of living index-numbers.

(2) Certain budget collections, where we have had access to the detailed individual returns, are studied more minutely so as to describe the relation of individual to average expenditure.

(3) The theoretical analysis of the distribution of expenditure is developed, and the theoretical formulæ established are compared with the empirical results already obtained.

In a logical order, (3) should perhaps come first. But the results of the first two studies are true as actual statements of behaviour and are of importance whatever is their theoretical background. The present order is adopted, moreover, in the hope that the treatment of (1) and (2) is intelligible to many who will be repelled by the more severely mathematical form in which (3) is necessarily cast.

The study is "econometrical" in the sense that it attempts to apply measurement to economic actions. It is well to distinguish it from the more familiar econometric analyses of the demand for goods. These usually consider the variations in the amounts demanded by a group of consumers in relation to price variations without reference to the differences between the incomes of the individuals who make up the group. Here, on the other hand, we do not consider price variations at all. We assume that prices are the same within the group to which the budget collection relates. We are obliged to ignore any difference in prices that in fact may exist owing to the different locations of the members of the group, or to their different incomes, and there is thus some want of precision in our results, especially when the group investigated is at all heterogeneous. In (3) it should be noticed that the final formulæ assume constant prices.

The independent variable we consider, therefore, is income—or rather total expended income. For one budget collection, prices are assumed constant. But

prices vary greatly from one budget collection to another and this fact prevents us making more than very rough comparisons between different groups, at different times or in different countries. Some tentative comparisons are, however, attempted.

Where the budgets are accurate, detailed and sufficiently numerous, many interesting investigations can be carried out and some of these would form a natural sequel to our study. We have limited ourselves, to a very large extent, to an attempt at the formulation of some general laws, expressible by simple formulæ, which may serve one of the main purposes of statistical methodology by replacing a complex of figures by a small number of exactly defined measurements. In so far as these formulæ are of general application, they should not only have immediate practical use but should also provide economists with fundamental material for theoretical analysis.

CHAPTER I

AVERAGE EXPENDITURES

1. Engel's Laws

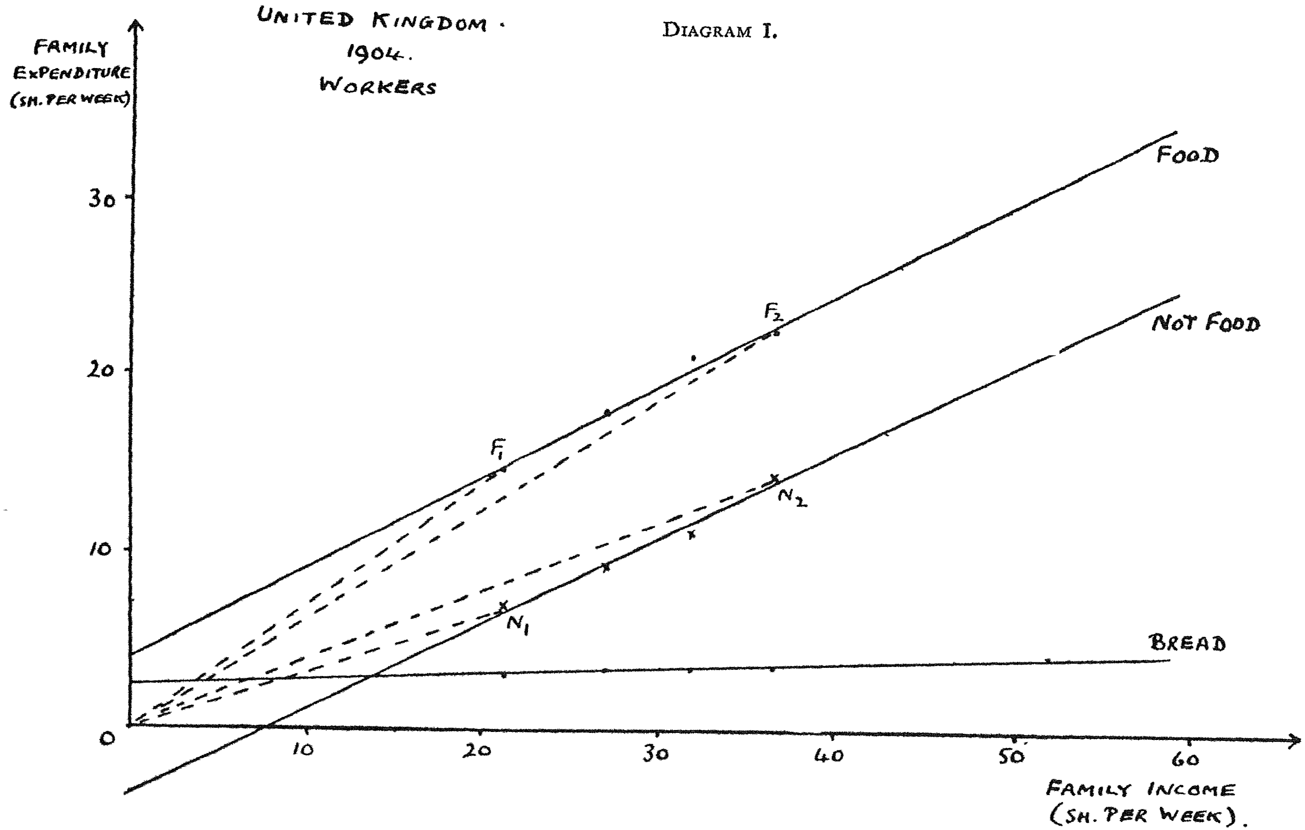
The results of collections of family budgets are usually summarised in a way shown by the Table below, which relates to the budget enquiry made by the Board of Trade in 1904. Bread is here selected as a typical food item; the full results would give average expenditures on many such items.

Working-class Budgets in the United Kingdom, 1904

Range of Weekly Income. sh.	Average Income. sh.	Average Expenditure on			Percentage of Income on		
		Food. sh.	Not-food. sh.	Bread. sh.	Food.	Not-food.	Bread.
Under 25 .	21·4	14·4	7·0	3·0	67	33	14·3
25 to 30 .	27·0	17·8	9·2	3·3	66	34	12·3
30 to 35 .	31·9	20·8	11·1	3·3	65	35	10·3
35 to 40 .	36·5	22·3	14·2	3·4	61	39	9·2
40 and over	52·0	29·7	22·3	4·3	57	43	8·3

We proceed to examine these expenditures to ascertain whether they are consistent with any simple formula. The relations of average expenditures to average income are shown graphically in Diagram I, where the plotted points represent the data on the ordinary scheme of rectangular axes. It is evident that the three sets of markings lie approximately on three straight lines and lines are ruled in the diagram to pass through or near the points marked.

DIAGRAM I.



UNITED KINGDOM.
1904.
WORKERS
FAMILY
EXPENDITURE
(SH. PER WEEK)
30
20
10
0
10
20
30
40
50
60
FAMILY INCOME
(SH. PER WEEK).
FOOD
NOT FOOD
BREAD
F_1
F_2
N_1
N_2

Rectilinear relations of this kind are to be expected if we are dealing with a group of families with the same complex of desires or preferences and the same needs, if the families differ only in respect of income and if the formula

$$y = ke + c$$

applies throughout the income range.[1] But we can formulate and examine the results objectively without any theoretical basis.

The formula suggested by these figures leads to a precise statement of what is known as Engel's Law. This law is to the effect that, as income increases, the expenditures on different items of the budget have changing proportions and that the proportions devoted to the more urgent needs (such as food) decrease while those devoted to luxuries and semi-luxuries increase.

If the rectilinear relationship between expenditure and income suggested by the figures above does in fact hold, Engel's Law can be reformulated more precisely as follows:

In a homogeneous group of families differing only in respect of income, the excess over (or defect from) the average of expenditures on any budget item bears a constant proportion to the excess over (or defect from) the average income. In the case of some goods, which may be described as *necessaries*, this rule results in a diminishing proportion of expenditure as income rises. In the case of other goods, which may be described as *luxuries*, the proportion of expenditure rises as income rises.

[1] Suppose that the expenditure y_t of the tth family on any item is related to the income e_t of the family by the formula

$$y_t = ke_t + c$$

where k and c are the same for all families. Taking the average for any n families and writing $\bar{y}$ and $\bar{e}$ for the average expenditure and income, we have

$$n\bar{y} = \Sigma y_t = k\Sigma e_t + nc = kn\bar{e} + nc$$

i.e.

$$\bar{y} = k\bar{e} + c$$

The same result is obtained for the averages if y_t differs from $(ke_t + c)$ by an amount v_t provided that the sum of the v_t's is zero over the range averaged in the group.

The test of the applicability of this expression of Engel's Law is the closeness of the fit of the points representing the data to straight lines.

In Diagram I, the straight line for food expenditure cuts the vertical axis through zero income at a point above zero. Consequently lines such as OF_1 and OF_2 show a decreasing gradient as income increases and a diminishing proportion of income is devoted to food. A similar statement is true of the straight line representing expenditure on bread. On the other hand, the straight line representing non-food expenditure cuts the vertical axis negatively, lines such as ON_1 and ON_2 show an increasing gradient and the proportion of income devoted to non-food items increases with income.

In general, it can be stated that the two classes of items, necessaries and luxuries, correspond to positive and negative values respectively of the constant c in the expenditure formula written above.

Theory, however, leads us to expect the rectilinear relation to hold only if family needs are similar or, perhaps, if needs vary in some regular way with income. In the 1904 budget enquiry, the number of children living at home increases with income, partly because the children themselves contribute to, or provide part of, the income. The aberrations of the points which relate to the top range of incomes in Diagram I are probably due to the larger number of children in the families in this range. In the sequel we make allowance for this variation of needs whenever the data are sufficient to make it possible.

If the rectilinear formula is valid for homogeneous groups of families, then departure from it would indicate that we are passing from one group to another. If budgets from working-class and from middle-class families were included in one collection, for example, the points corresponding to the latter might be found to lie on a different straight line from those of the former.

But theory would lead us to regard the linear law as only a first approximation to a regular curve. If the data,

when represented graphically, show a regular curvature, the statement of Engel's Law would need elaboration. But we have found that, in the budget collections examined, the simple statement is usually sufficient except perhaps at the upper extreme. The examination of the curvature, if it exists, would need a very large collection and more income classes than are usually given. Some indication of curvature, in the cases of certain items, will be found in the sequel. Regular change of direction, which is to be expected in a homogeneous group with a non-linear expenditure relation, is quite a different matter from a sudden shifting from one straight line to another. The latter is an indication of the presence of two groups with different complexes of preferences or, possibly, with different market conditions.

As we pass from one country to another, or from one date to another, we are likely to be in the presence of different groups with different modes of expenditure and changed desires, quite apart from the difference of real incomes. If the relative prices of different goods were the same in the two situations (whatever the actual scale of prices), we could examine the nature of this variation by considering the changes in the constants k and c from one situation to the other. Since relative prices do change, however, we can only go a little way in this direction. Where bread is cheap and meat dear, for example, a family with given desires will allot expenditure of the same income in a different way from that when bread is dear and meat cheap.

The diagrams of the present chapter, together with the tabled values of the constants, will show these differences whether they are due to changed desires, to changed prices or to both. The scales in the diagrams are so chosen that their gradients have the same significance throughout.

2. The Scale of Urgency

The validity of the linear expenditure law can be observed from a graphical presentation of the data, and

the relative gradients and positions of the lines representing different expenditures can then be interpreted as expressing the relative urgency of needs. When we come to actual measurement, however, it is necessary to develop the formulæ algebraically.

Taking total expenditure (i.e. total expended income) as the independent variable to be measured horizontally in the diagrams, write f for the expenditure of any family on some defined item, e.g. on food as a whole or on some specific food item. The fundamental relation is

$$f = ke + c \quad . \quad . \quad . \quad (1)$$

where k and c are constants to be found from the data.

Our method has usually been to draw a careful diagram, to rule a line which appears to lie as close as possible to the plotted points (particularly in the central region) and then to choose two of the observations which fit closely to the line, computing the values of k and c from these two points. If the law is perfectly obeyed, of course, any two observations suffice. In some cases, the constants have been determined by a least squares formula, but this has not always been as satisfactory as more empirical methods. In any case, we have only aimed at approximate results since the nature of the material has never allowed of great precision. In the following chapter, as will be indicated, more rigid methods are applied.

Write $\bar{e}$ and $\bar{f}$ for the average total expenditure and average expenditure on the particular item considered, as computed for the whole group of families comprising the budget data. Then the point representing $\bar{e}$ and $\bar{f}$ should lie on the line (1). Hence

$$\bar{f} = k\bar{e} + c$$

and subtraction from (1) gives

$$f = ke + (\bar{f} - k\bar{e}) \qquad (2)$$

Write $\qquad \bar{w} = \bar{f}/\bar{e}$

so that $\bar{w}$ is the proportion that the special expenditure

bears to the average total expenditure. Hence, the equation (2) can be written

$$f = ke + (\bar{w} - k)\bar{e}. \qquad . \qquad . \qquad (3)$$

Evidently

$$c = (\bar{w} - k)\bar{e} \qquad . \qquad . \qquad . \qquad (4)$$

By this adaptation, we have obtained constants k and $\bar{w}$ which are completely independent of scale and units; their values are unaffected whether expenditures are for a week, month or year, or expressed in shillings, pounds, marks or any other currency unit. With the addition of a statement of the value of $\bar{e}$ in the appropriate currency units, these two constants are sufficient for all further analysis, in so far as the linear law (1) is applicable. It should be noticed that the constant c, as given by (4), is dependent on the currency units and on the length of period to which the expenditures refer.

From (3) we obtain at once

$$f/e = k + (\bar{w} - k)\bar{e}/e \qquad . \qquad . \qquad (5)$$

The expression (5) represents the proportion of total expenditure devoted *by any family* to the special item considered. If $\bar{w}$ is greater than k, this proportion decreases as total expenditure increases. On the other hand, if $\bar{w}$ is less than k, the proportion increases with total expenditure. An item has been defined as a necessity if c is positive; it is now seen that an equivalent definition is provided by the inequality $\bar{w} > k$. Similarly, an item is a luxury if c is negative and $\bar{w} < k$.

If all items in the family budget are included, the total of the different $\bar{w}$'s is necessarily unity. It is easily seen that the total of the k's, for all items, is also unity.[1] The total of the $(\bar{w} - k)$'s is thus zero, as is the total of the c's.

In the ordinary statements of the results of budget enquiries, the proportion of expenditure on any item, our $\bar{w}$, has a quite familiar meaning and it is usually tabulated as a percentage. It is used as the basis of cost of living index-numbers. On the other hand, k is not

[1] See Chap. III, § 4, below.

considered in such enquiries and needs more interpretation. When computed from a detailed collection of budgets, as in Chapter II below, k is the co-efficient of regression familiar to statisticians. Then, as also when it is obtained from averaged budgets, it measures the gradient of the straight line through the plotted points which represent the data.

It is more useful, however, to use, not k and $\bar{w}$ alone, but also either $(\bar{w} - k)$ or $k/\bar{w}$. As already shown, the sign of $(\bar{w} - k)$ determines whether an item is a necessary or a luxury. It can now be added that the values of $(\bar{w} - k)$ for the various items of the budget can be used to determine an order of urgency amongst the items. (The value of $(\bar{w} - k)$ for any item is by (4) the value of the constant c as a proportion of the average total expenditure.) If $(\bar{w} - k)$ is positive, the larger its value the more rapidly does the proportion of expenditure on the item decrease as total expenditure increases. Conversely, if $(\bar{w} - k)$ is negative, the larger its numerical value the more rapidly does the proportion increase with total expenditure. The order of the values of $(\bar{w} - k)$ from large positive to large negative values thus fixes what may be termed the order of urgency of the items in the budget.[1] This provides the basis of the ordering of the items in the group of diagrams given in § 7 of the present chapter.

The ratio $k/\bar{w}$ serves a very similar purpose. It measures what is known as the *income elasticity of demand* for the special item at the average level of total expenditure. That is, it is the ratio of the proportionate increase of expenditure on the item to the proportionate increase of total expenditure (which we have here taken instead of income) as we go a step up the scale of total expenditure

[1] These statements follow from (5). We have

$$\frac{d}{de}\left(\frac{f}{e}\right) = -\frac{(\bar{w} - k)\bar{e}}{e^2}$$

The value of $(\bar{w} - k)$ thus determines the rate at which the proportion of expenditure on the item increases or decreases, as total expenditure increases from any level. The order of urgency is defined by the order of these rates.

from the average level. For example, if $k/\bar{w}$ is 1·2, then a 1 per cent increase of total expenditure from the average corresponds approximately to an increase of 1·2 per cent in the expenditure on the item. In the sequel we denote the ratio $k/\bar{w}$ by the single symbol $\bar{\eta}$.[1]

In rare cases, k (and so $\bar{\eta}$) is negative. This is the case when, with a larger total expenditure, less is bought of a good for which there is a preferable substitute. k and $\bar{\eta}$ are zero when the same amount of a good is bought whatever the total expenditure. As $\bar{\eta}$ increases to unity, k becomes larger relative to $\bar{w}$ and we have diminishing pressure of necessity, while as $\bar{\eta}$ rises still further the items in question become more definitely luxurious. The value unity of $\bar{\eta}$ divides the class of luxuries from the class of necessaries. The reverse order of values of $\bar{\eta}$ is thus also an indication of the order of urgency. It is, perhaps, a less good criterion than the previous one since $\bar{\eta}$ is defined at the average expenditure whereas the first criterion depends on the constant c. Our object, however, is never more than an indication of an approximate order of urgency.

The value of $\bar{w}$ for any item, unlike the values of k and c, evidently depends on the range and grouping of incomes included in the particular budget collection examined. The value of $\bar{\eta}$, though again independent of units, is also dependent on the same accident of choice.

[1] The concept of income elasticity of demand is discussed in the Mathematical Appendix, § 3, below. It is defined

$$\eta = \frac{e}{f}\frac{df}{de} = \frac{ke}{ke + (\bar{w} - k)\bar{e}}$$

Hence, $\bar{\eta} = k/\bar{w}$ when $e = \bar{e}$, and it follows that

$$\frac{\bar{\eta}}{\eta} = \frac{ke + (\bar{w} - k)\bar{e}}{\bar{w}e} = 1 + (1 - \bar{\eta})\left(\frac{\bar{e}}{e} - 1\right)$$

If $\bar{\eta} = 1$, then $\eta = 1$ at all expenditure levels. Further, if $\bar{\eta} < 1$, η increases and tends to a limiting value unity as e increases, i.e. η increases but remains less than unity at all expenditure levels. If $\bar{\eta} > 1$, η decreases as e increases but remains greater than unity at all expenditure levels. Hence, the order of the $\bar{\eta}$'s for various items can also indicate the order of the η's at any expenditure level, at least not far removed from the average. The spread of the η's about unity decreases as expenditure increases.

If, however, the selection of budgets is typical of a defined class, $\bar{\eta}$ has a definite significance. This is reinforced by the fact that the income elasticity of demand changes slowly as we move from the average level of income.

Our definition of an order of urgency, it may be objected, implies that the expenditure on a luxury item, with a negative value of c in the formula (1), becomes negative for low incomes. In other words, it appears that a family with a small expenditure spends more than the whole of the expenditure on necessaries and a negative sum on luxuries. This is, perhaps, not as ridiculous as it appears. But, in any case, it has been pointed out that the linear law (1) is only a first approximation. The approximation is only a close one over the central range of incomes and may easily cease to apply at low or high incomes. The order of urgency is to be considered as limited to the central range of incomes to which the linear law (1) applies.

The above analysis, though expressed in terms of expenditures on specific items, is clearly applicable also to expenditures on groups of items. We obtain relations and measurements of the same form whether we take items such as beef, mutton, pork and so on separately, or whether we combine them into one comprehensive meat group. Suppose that the linear law (1) holds for each separate item:

Beef . . . $f_1 = k_1 e + c_1$
Mutton . . . $f_2 = k_2 e + c_2$
Pork . . . $f_3 = k_3 e + c_3$ . . .

By addition, we have

$$F = Ke + C$$

where F is the expenditure of a family on the meat group as a whole and where K and C are constants given as the sum of k_1, k_2, k_3, . . . and of c_1, c_2, c_3, . . . respectively. A linear relation thus holds for the expenditure on the meat group. Further, if $\bar{W}$ is the sum of the $\bar{w}$'s appropriate to the separate items, then $\bar{W}$ is the proportion of the average total expenditure devoted to the meat

group. It follows, finally, that the ratio $K/\bar{W}$ measures the income elasticity of demand for the meat group as a whole.

3. Cost of Living Index-numbers

Before exhibiting the results of our analysis of actual budget collections, it will be interesting to show the applicability of Engel's Law, as restated above, to cost of living index-numbers.

It is commonly, and correctly, said that an index of the cost of living based on average expenditure is only applicable to families whose expenditures are near the average. Further, quite apart from expenditure, if an index is appropriate to one social class, then it is incorrect to apply it to another.

We are now in a position to test the importance of these criticisms. In the first place, if we have budgets of working-class family expenditure and also budgets of middle-class expenditure, then a graphical exhibition of the two sets of data, and a study of the corresponding values of $\bar{\varpi}$, k and c, will enable us to determine how far a formula for expenditure on any item obtained from one class of budgets is applicable to the other class. In so far as a single formula is applicable to all families, we can then proceed as follows.

Suppose that $\bar{e}$ is the average total expenditure of a group of families and $\bar{\varpi}_t$ the proportion of $\bar{e}$ devoted to the tth item. Let r_t denote the ratio of the price of the tth item in the current year to its price in a fixed base year. Then the cost of living index-number at the average expenditure is given by the formula

$$I = 100\Sigma(\bar{\varpi}_t . r_t)$$

where, of course, the sum of the weights $\bar{\varpi}_t$ is unity.

Write $$r_t = \frac{I}{100} + \rho_t$$

so that $$\Sigma(\bar{\varpi}_t . \rho_t) = 0$$

For any expenditure e, other than the average, our rectilinear formula gives, for the actual expenditure on the tth item,

$$e_t = k_t e + (\bar{w}_t - k_t)\bar{e}$$

where $$\Sigma k_t = 1.$$

The correct index-number for the expenditure e is then

$$\begin{aligned} I_e &= 100\Sigma\left(\frac{e_t}{e}r_t\right) \\ &= I\Sigma\left(\frac{e_t}{e}\right) + 100\Sigma\left(\frac{e_t}{e}\rho_t\right) \\ &= I + 100\Sigma\left[\rho_t\left\{k_t + (\bar{w}_t - k_t)\frac{\bar{e}}{e}\right\}\right] \end{aligned}$$

since $\Sigma e_t = e$.

i.e. $$I_e = I + 100\left(1 - \frac{\bar{e}}{e}\right)\Sigma(\rho_t k_t).$$

The summation in this formula covers all the items in the family budget. It can be applied in the case where every item is priced separately but, to illustrate its use, we take the items in expenditure groups.

The figures exhibited in the following table are taken from the *Ministry of Labour Gazette* of January, 1935:

	$\bar{w}$	r	ρ	k	ρk
Food	·60	1·25	— ·18	·36	— ·0648
Rent	·16	1·56	·13	·06	·0078
Clothing	·12	1·875	·445	·18	·0801
Fuel and Light . .	·08	1·725	·295	·016	·0047
Miscellaneous . .	·04	1·725	·295	·384	·1133
Total . .	1·00			1·00	·1411

The first column of figures gives the actual weights used in the Ministry's cost of living index-number, i.e. the basic proportions of average expenditure. The second column gives the relative price changes from July, 1914,

to January, 1935. The third column is computed from the definition of ρ given above; each value of ρ shows the difference between the particular price change and the average index-number $\bar{I}/_{100} = 1{\cdot}43$. The latter is calculated by applying the weights $\bar{w}$ to the r's multiplied by 100. The fourth column gives the values of k for the groups of items and, in default of appropriate British statistics, it is taken from the equations obtained for German working-class expenditure (1927–8).

The value of the index-number for any expenditure is then

$$I_e = 143 + 14{\cdot}11\left(1 - \frac{\bar{e}}{e}\right)$$

For example, if $e = 2\bar{e}$, then $I_e = 150$
and if $e = \frac{1}{2}\bar{e}$, then $I_e = 129$.

Thus, to allow for the purchase in January, 1935, of the same quantities and kinds of goods as in the base year, the income of an average working-class family (to which the $\bar{w}$'s are supposed to apply) would need to be raised to 43 per cent above its level in the base year. On the other hand, a clerk with twice the income would need 50 per cent increase in income as compared with the base year so far as expenditure on the goods included is concerned. A workman with less than the average income would need a smaller rise. These results would, of course, be somewhat modified if the computation was carried out in full detail.

It can be computed quite easily that the values of the proportionate expenditures of the clerk's family in the base year are to be taken as ·48 for Food, ·11 for Rent, ·15 for Clothing, ·048 for Fuel and ·212 for Miscellaneous. The proportions in January, 1935, become ·40, ·11, ·19, ·055, and ·245 respectively. These figures are obtained from the linear expenditure relation given above.

This calculation is given only as an illustration of the method; we have not the information necessary to make an exact estimate for Great Britain on this basis.

4. Expenditure in Relation to Needs

The theoretical analysis of the distribution of expenditure postulates a group of families with the same needs and complex of preferences and with differing incomes. We must expect, therefore, to obtain the linear expenditure relation suggested by theory only if the needs and composition of the families investigated are reasonably similar.

The condition of similar needs is, at best, fulfilled only approximately in actual budget collections. The families usually consist of a man and wife with a variable number of children, some of whom may be earning and contributing to the family income. There may be also a number of childless families and families with additional adults or with lodgers. The needs of a family vary widely with its composition as regards age, sex and earning power. We must, therefore, find some way of eliminating the effect of the variation of needs, whenever the data are sufficient for the purpose, if we are to exhibit a definite relation between the expenditure on any item and the total expenditure.

Expenditures can be divided, broadly speaking, into two categories, those common to the family and those special to the individuals of the family. The first category includes rent, fuel, lighting and some miscellaneous items. In the other category, we have clothing and the various food items.

An obvious method of allowing for the variation of needs in the expenditures of the first category is by assuming a formula of the form

$$r = ke + bn + c$$

Here, r stands for (e.g.) rent expenditure, e for total expenditure as before and n is an index of the family size. In computing the value of n, we can either take each individual as a unit or count children as fractions on some scale related to age. For families with the same general composition, i.e. with the same value of n, the formula

reduces to the familiar linear relation between one expenditure and the total expenditure, the constant term in the relation being provided by the constant expression $(bn + c)$.

The co-efficients k, b and c are to be derived by the usual method of partial regression, using the data provided by the budget collection considered. If the formula applies rigidly to the families concerned, it gives the actual rent expenditure of any family for which the values of e and n are known. The constant k indicates the variation of rent expenditure as a proportion of the variation in the income of families of given composition.

One of the results obtained from the English middle-class budgets (1926), and utilised in the following chapter, is expressed

$$r = \cdot 109e - 3\cdot 78n + 17\cdot 4$$

where r is weekly rent (average 27·5 shillings), e is total weekly expenditure (average 192 shillings) and n is the number of equivalent adults (average 2·86). The weekly rent paid by any family, on this formula, is thus 17·4 shillings, *plus* a little more than one-tenth of the family income, *minus* 3·78 shillings for every equivalent adult. The amount allotted to rent for a given income decreases, in this group, as the size of the family increases. On the other hand, for families of a given size, the rate of increase of rent expenditure is roughly one-tenth of that of total expenditure. The presence of the fixed sum of 17·4 shillings shows that the constant in the linear relation of rent to income is not entirely, or even mainly, dependent on the number of persons in the family.

For the second category of expenditures, those special to the individual, it is clearly more appropriate to divide each expenditure by the number of persons. This number should be, not the actual number of individuals, but a number based on a scale of needs in which allowance is made for age and sex. Many such scales of equivalence have been devised. The usual method is to take

an adult man as a unit and to count women and children as appropriate fractions. It is not necessary to discuss the scales in detail here since the various scales appear to give nearly the same results in the present study and since, in any case, the scales have usually been chosen and used in the reports of the budget collections on which we draw.

In considering family expenditure on food items, therefore, we proceed as follows. We now take, not the total expenditure of a family, but the total food expenditure as the independent variable. For each family, the number of equivalent adults is computed according to some chosen scale of equivalence and all food expenditures are divided by this number. The linear relation between expenditure on any food item (e.g. the meat group) and the total food expenditure is then written

$$m = kf + c$$

where m and f are the expenditures on meat and on all food per equivalent adult. The constants k and c, and the measurements $\bar{\varpi}$ and $\bar{\eta}$ derived from them, have meanings exactly analogous to those already discussed. The only difference is that the expenditures from which they are derived are not per family but per equivalent adult.

In the following chapter, we are able to eliminate the effect of total expenditure and of needs in the case of each family and so to examine the nature of the residual variations of expenditure. In the present chapter, however, we only deal with the relations of average expenditures. But, for reasons already explained,[1] we expect average expenditures to satisfy linear relations if the expenditures of the separate families do so, or if the variations of family expenditures are regularly distributed.

The applicability of this modified linear law is tested as before by plotting the average expenditures per equivalent adult, as given by the data, on a graph and by fitting

[1] See footnote, p. 7 above.

straight lines to the plotted points.[1] If the approximations to linearity are close, then the restatement of Engel's Law given above can be expressed in a more developed form in terms of expenditures per equivalent adult, a form which applies when the families are not homogeneous in respect of needs.

We can, as before, consider items such as beef, mutton or pork separately or merge them into a meat group. The constants k and c for a group are again sums of the corresponding constants for the items separately. If a complete list of items, or of groups of items, is taken, the sum of the k's is unity and the sum of the c's zero.

Another method of approach to the problem of accounting for needs has also been tested. Suppose that a standard food diet for an adult man has been fixed in which account is taken of such considerations as calorific content.[2] The cost of the diet at the ruling prices is then found to be d shillings per week. Then suppose that the diet is reduced for other persons according to the following scale of equivalence:

Adults, 100; Schoolchildren, 60; Younger children, 30; where, for simplicity, no account is taken of sex differences. The food needs of a family comprising n_1 adults, n_2 schoolchildren and n_3 younger children now costs

$$(n_1 + \cdot 6n_2 + \cdot 3n_3)d$$

shillings per week. On the assumption that this cost is always incurred irrespective of income and that the remainder of the food expenditure depends solely on

[1] In the averaged budgets used in this chapter, we have not always been able to use data in which the expenditures per equivalent adult have been calculated for each family. In some cases there were only statements of the average number of equivalent adults or of children in the families within each income grade. The average expenditures on various items divided by the average number of equivalent adults in the appropriate grades was then used instead of the expenditure per equivalent adult averaged over all families. In such cases, we cannot expect exact conformity of expenditures to the above or to any other simple formula.

[2] See, for example, Rowntree, *Poverty: A Study of Town Life*.

income, the total food expenditure of a family of the above composition is then given by the formula

$$f = ke + c + (n_1 + \cdot 6n_2 + \cdot 3n_3)d$$

where e is the weekly total expenditure in shillings.

The Liverpool working-class budgets (1929) were used to test a formula of this kind. The data were allowed to determine the co-efficients of n_1, n_2 and n_3 and the results were compared with the formula as written above. By the usual method of partial regression, it was found that

$$f = \cdot 33e + \cdot 44 + 2\cdot 8n_1 + \cdot 12n_2 + \cdot 36n_3$$

Hence, food expenditure, at least in this group of families, is not determined by needs in the sense here described. The important relation is clearly between food expenditure and income.

If we pay no attention to income, the result is

$$f = 6\cdot 3n_1 + \cdot 14n_2 - 2\cdot 6n_3 + 10\cdot 5$$

and this shows that the weekly expenditure on food actually decreases when the number of young children in the family increases. The number of such children is larger in the poor families of this group than in those better off.

5. Numerical Illustrations of the Linear Expenditure Law

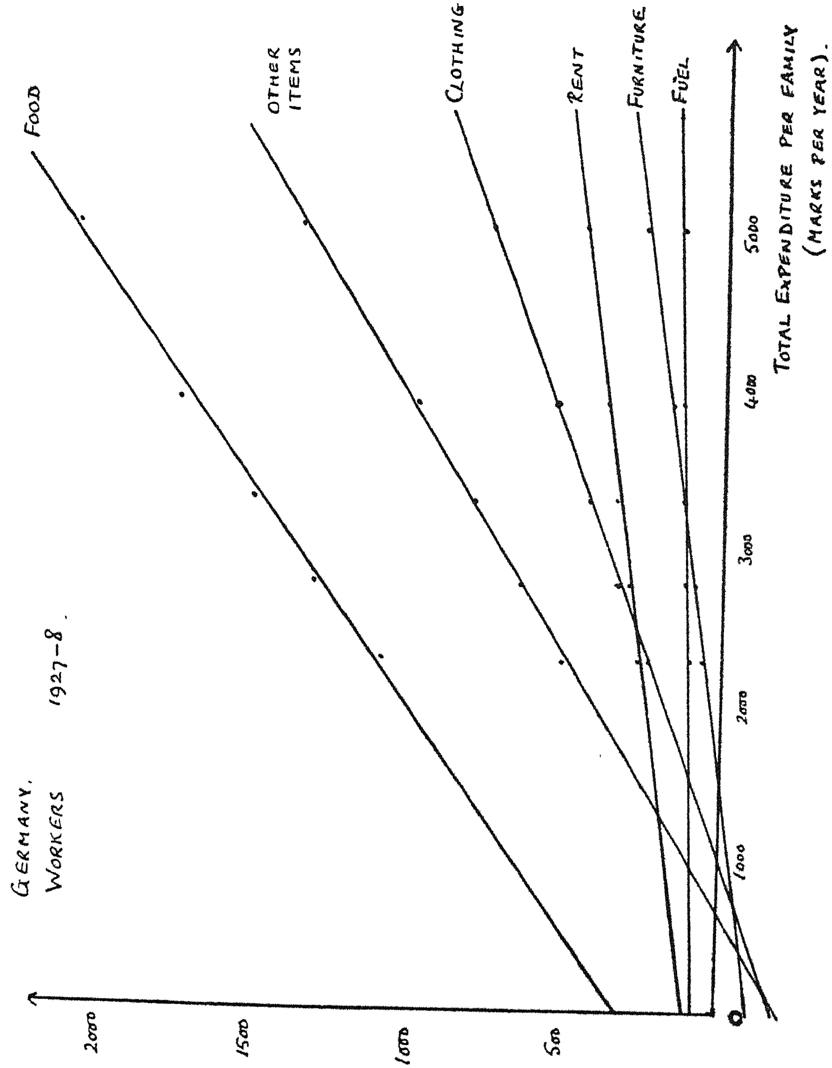
We have made use of a wide range of budget collections and a list of these collections, together with a brief description of their nature, is given at the end of this chapter. It has not been our intention to attempt a general or complete account of what is known about family expenditure, still less to discuss all the various analyses that can be usefully applied to budgetary statistics. Our main objects have been to test the validity of the expenditure formulæ developed above and to illustrate the use of the concepts derived from the formulæ. Some interesting comparative results can, however, be obtained.

The selection of budgets for our purpose has been determined by the nature of the data available and not with any intention of proving any *à priori* hypothesis. The budgets vary in date, in country and in completeness of information. The definitions and classifications of items into groups also vary, but we have not thought it necessary to discuss them in more detail than is given in the references below.

The six diagrams II–VII illustrate the applicability of the linear expenditure formula to certain budget collections. The data, in every case, relate to the average expenditures in various ranges of income (or of total expenditure) as defined and given in the original reports. The diagrams show these average expenditures plotted against the corresponding total expenditures according to the graphical method already described.

The first two diagrams take family expenditure as a whole and no allowance is made for the differences between the number and ages of the persons composing the families. It is seen that straight lines represent the expenditures of the German workers adequately in all cases, except that there is some slight evidence of curvilinearity in the case of food expenditure. The same is true of the American data except for some divergence in the higher income groups in the cases of rent and clothing. The data here relate to budgets collected from all parts of the U.S.A. and from families with a very wide range of incomes and of various social positions. We have, therefore, a good test of the extended application of the linear formula. The implication, if the graphic representation shows a rectilinear relation, is that tastes or preferences are common over a wide range of families. It is possible that the divergence in the case of rent or clothing is due to a definite change in the tastes of the families as we proceed to the higher income groups; such a change is, at least, to be looked for in a comprehensive collection of this kind. But the number of families in the upper groups is relatively small and the top income range is of undefined extent. We can only conclude that

DIAGRAM II.


GERMANY.
WORKERS
1927–8.
FOOD
OTHER ITEMS
CLOTHING
RENT
FURNITURE
FUEL
TOTAL EXPENDITURE PER FAMILY (MARKS PER YEAR).
0
1000
2000
3000
4000
5000
500
1000
1500
2000

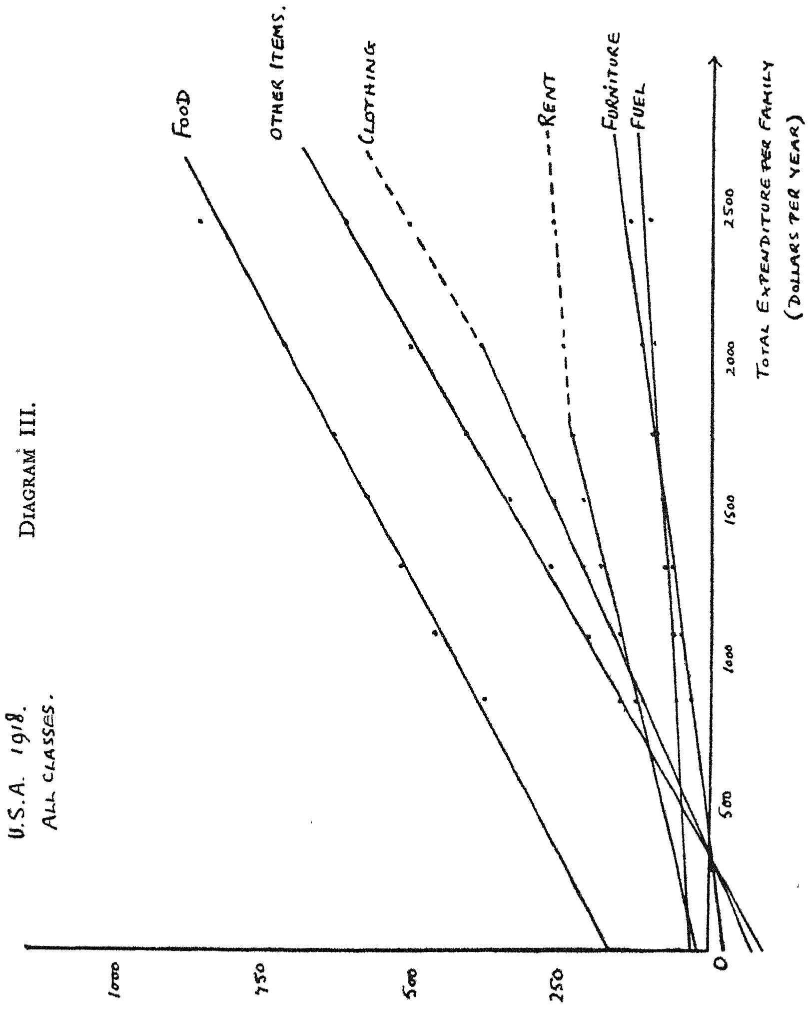

U.S.A. 1918.
ALL CLASSES.
FOOD
OTHER ITEMS.
CLOTHING
RENT
FURNITURE
FUEL
TOTAL EXPENDITURE PER FAMILY
(DOLLARS PER YEAR)
0
500
1000
1500
2000
2500
250
500
750
1000

DIAGRAM III.

DIAGRAM IV.

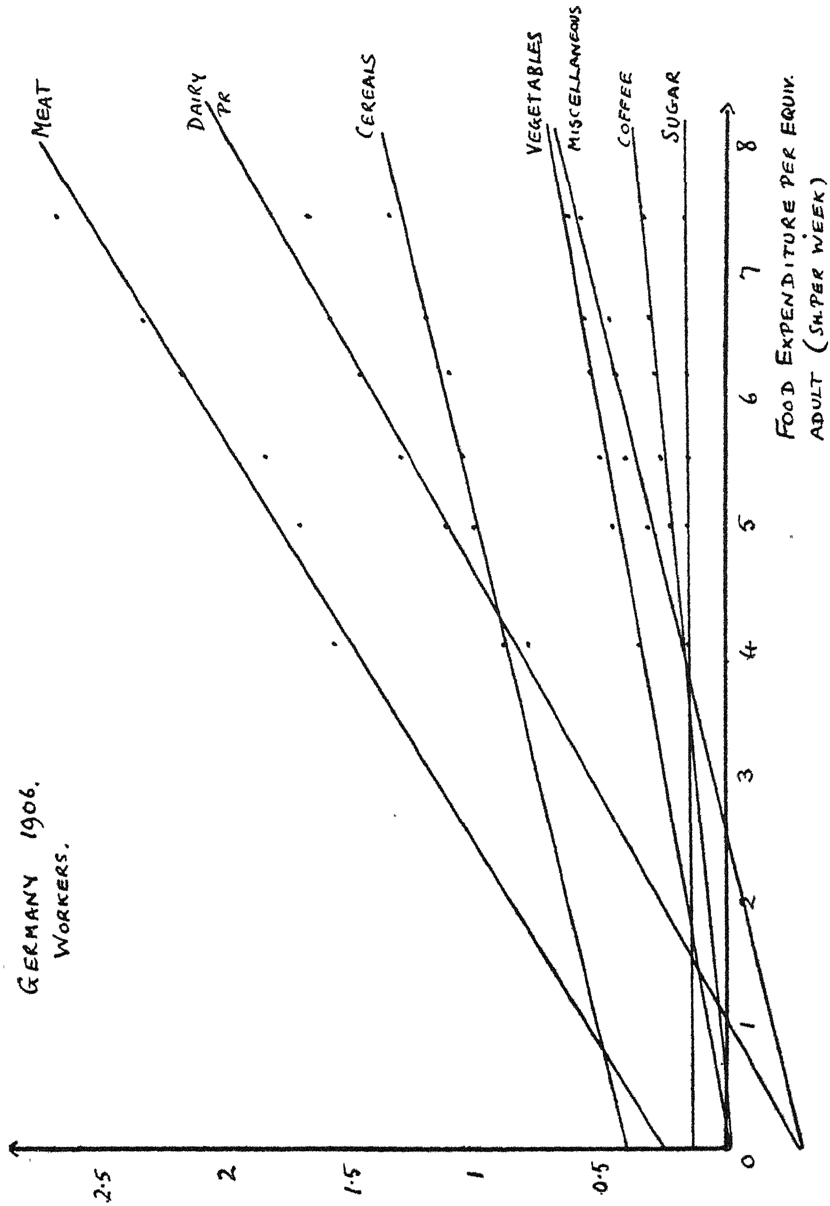
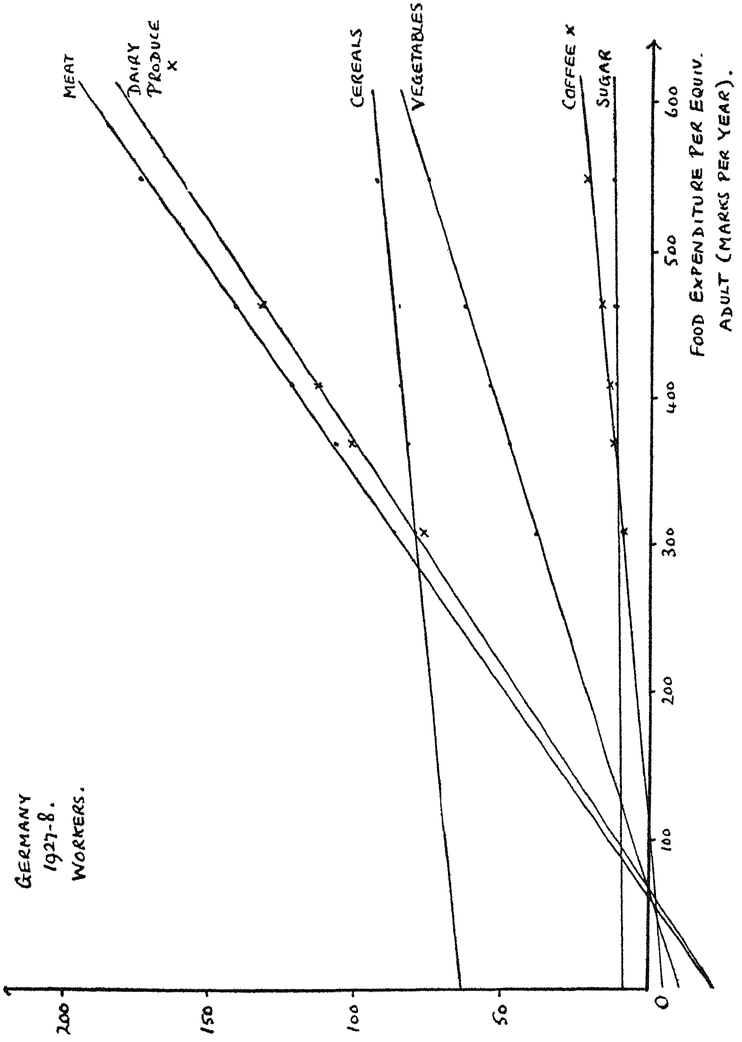

GERMANY 1906.
WORKERS.
MEAT
DAIRY PR
CEREALS
VEGETABLES
MISCELLANEOUS
COFFEE
SUGAR
FOOD EXPENDITURE PER EQUIV. ADULT (SH. PER WEEK)
0
1
2
3
4
5
6
7
8
0·5
1
1·5
2
2·5

DIAGRAM V.



(THE MISCELLANEOUS LINE IS PARALLEL TO, AND A LITTLE BELOW THAT OF COFFEE)

DIAGRAM VI.

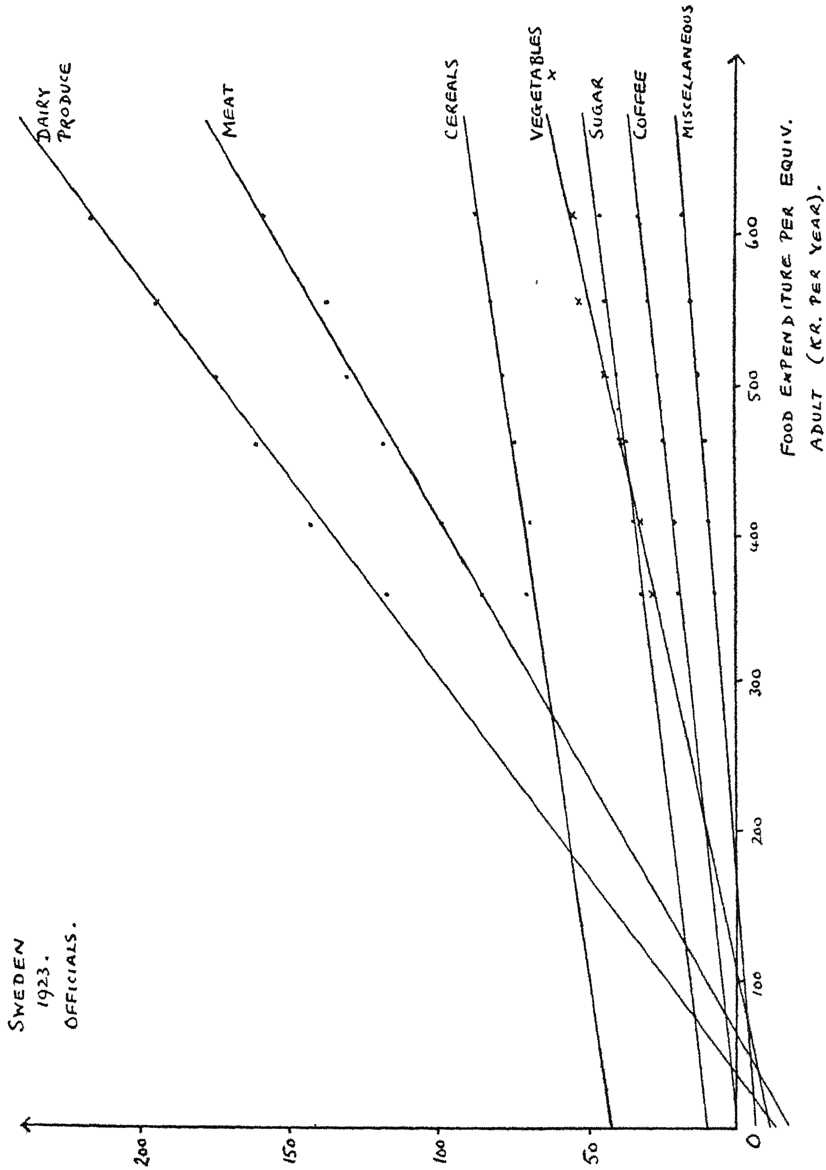
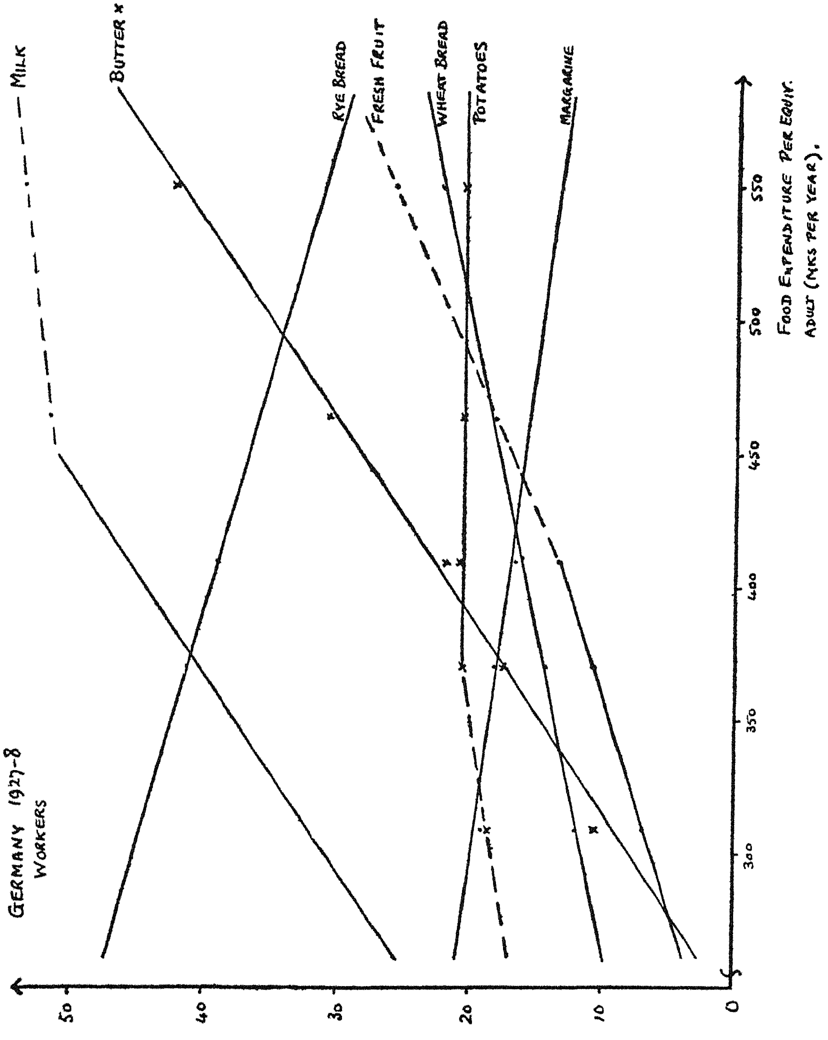

SWEDEN
1923.
OFFICIALS.
DAIRY PRODUCE
MEAT
CEREALS
VEGETABLES
SUGAR
COFFEE
MISCELLANEOUS
FOOD EXPENDITURE PER EQUIV. ADULT (KR. PER YEAR).
0
100
200
300
400
500
600
50
100
150
200

DIAGRAM VII.


GERMANY 1927-8
WORKERS
MILK
BUTTER x
RYE BREAD
FRESH FRUIT
WHEAT BREAD
POTATOES
MARGARINE
FOOD EXPENDITURE PER EQUIV.
ADULT (MKS PER YEAR).
300
350
400
450
500
550
0
10
20
30
40
50

more detailed examination is needed to determine whether expenditure lines of different gradients are obtained when the families are divided into more homogeneous groups.

The next three diagrams relate to expenditure on food and on the main food groups and allowance is now made for differing family composition by taking expenditures per equivalent adult. Diagrams IV and V are drawn from budget data of German workers at different dates. At the earlier date, the reduction to equivalent adults has been made on the rough basis of counting each child as half an adult, the process being applied to the average family in each income range. This can only give approximate results and it may be for this reason that the meat and dairy expenditures diverge from linearity. On the other hand, it may be that these expenditures show a curvilinear relation to total expenditure. In the diagram showing the data at the later date there are only trifling aberrations of the plotted points from straight lines. Diagram VI shows expenditures in a group of Swedish "Officials," a group which includes civil servants and other workers in the permanent government services. The divergences from straight lines are again not considerable and we can assume, therefore, that the class is fairly homogeneous, on the average, in respect of tastes or preferences.

In Diagram VII is shown the variation of expenditure on food of the same group of German workers as Diagram V. Specific items of food are now taken instead of the food groups. The divergences from linearity are seen to be quite small in general, consideration being given to the fact that greater relative variability is to be expected when the proportion of expenditure is smaller. There are, however, certain cases which merit special attention. There is some evidence of a curvilinear relation in the case of milk expenditure. On the other hand, there may be two distinct expenditure lines, one for the lower and one for the higher income groups, and the lines shown in the Diagram are intended to illustrate this possibility. In the case of expenditure on potatoes, the lowest income

group diverges from the expenditure line clearly established by the other income groups. It is possible that a definite change of tastes occurs here. Finally, in the case of fresh fruit, two distinct expenditure lines can be drawn and considerable evidence of a change of tastes occurring about the middle of the income range is provided by the data. It is not safe to draw any definite conclusions, however, without further examination of the budgets. The doubtful case of milk expenditure, in particular, requires further consideration.

The averaged budget collections illustrated by these diagrams have been selected from a much larger number of collections, all of which tend to show that the hypothesis of the straight line expenditure relation is not unjustified. The graphical method of testing this hypothesis is easily carried out and any statistician can make further experiments. Space does not permit the representation of other data here.

6. Some Comparative Results

In Tables A, B and C we exhibit the results obtained from the large number of budget collections at our disposal. The tables give the values of $\bar{e}$, the average total expenditure, and for the various items of expenditure enumerated the values of $\bar{w}$, the average proportion of total expenditure ; of k, the gradient of the line representing expenditures at different total expenditure levels; of $\bar{\eta}$, the income elasticity of demand at the average expenditure level. To save space, the values of the constant c of the expenditure line are not tabled; they can be computed, for each item, from the formula

$$c = (\bar{w} - k)\bar{e}$$

Table A relates to the main classes of family expenditure and no correction has been made for the varying size and composition of families. In the construction of Table B, the expenditures on the main classes of food expenditure have been related to the total food expenditure and all expenditures have been divided by the num-

TABLE A. FAMILY

Family Expenditures on the Main Groups of Items

	United Kingdom, 1914. Workers.	Liverpool, 1929. Workers.	English Towns, 1926. Clerks.	L.S.E., 1932. All Classes.	Belgium, 1928–9. Workers.	Amsterdam, 1923–4. All Classes.	Germany, 1927–8. Workers.	Germany, 1927–8. Salaried Employees.	Germany, 1927–8. Officials.	Copenhagen, 1922. All Classes.
Number of Budgets . . .	—	154	112	123	116	212	896	542	498	102
Average Total Expenditure per annum	—	£132	£482	£436	21,156 Francs	3,133 Florins	3,165 Marks	4,352 Marks	4,925 Marks	6,719 Kr.
$\bar{w}$ Food	·60	·47		·26*	·57	·37	·45	·35	·35	·29
Rent	·16	·21	·60	·14	·06	·18	·10	·11	·12	·13
Clothing	·12	·08		·10	·15	·10	·13	·125	·14	·12
Fuel and Light . . .	·08	·09	·07	·04	·04	·05	·04	·035	·04	·05
Furniture	—	—	—	—	·03	—	·04	·06	·07	·06
Other Items	·04	·15	·33	·46	·16	·30	·24	·32	·28	·35
k Food		·42		·13*	·47	·20	·36	·27	·23	·09
Rent		·16	·47	·10	·03	·18	·06	·07	·12	·13
Clothing		·12		·09	·17	·15	·18	·15	·15	·11
Fuel and Light . . .		·04	·06	·02	·01	·02	·02	·02	·04	·04
Furniture		—	—	—	·025	—	·07	·10	·10	·08
Other Items		·26	·47	·66	·295	·45	·31	·39	·36	·55
$\bar{\eta}$ Food		0·9		0·5*	0·8	0·5	0·8	0·8	0·7	0·3
Rent		0·8	0·8	0·7	0·5	1·0	0·6	0·7	1·0	1·0
Clothing		1·4		1·0	1·1	1·5	1·4	1·2	1·1	0·9
Fuel and Light . . .		0·5	0·8	0·5	0·2	0·4	0·4	0·6	1·0	0·8
Furniture		—	—	—	0·7	—	1·8	1·8	1·4	1·3
Other items		1·7	1·5	1·5	1·8	1·5	1·3	1·3	1·3	1·6

NOTES.—*Total Expenditure* excludes taxes, savings and repayment of debts.
Food usually excludes alcoholic drinks and meals away from home.
Rent usually includes rates.
Furniture includes household linen and utensils.
Other Items usually includes washing and cleaning materials.

EXPENDITURE
in relation to Total Family Expenditure

Denmark, 1922. All Classes.	Oslo, 1912–13. All Classes.	Oslo and Bergen, 1918–19. All Classes.	Finland, 1920–1. Workers.	Basle.		Czechoslovakia.				Poland, 1929. Workers.	U.S.A.	
				1912. All Classes.	1921. All Classes.	1927. Workers.	1929. Workers.	1927. Officials.	1929. Officials.		1918. All Classes.	1928–9. Farmers.
141	85	75	437	78	64	79	262	122	291	84	12,096	269
5,610 Kr.	1,820 Kr.	5,610 Kr.	17,428 Marks	3,135 Francs	6,272 Francs	17,100 Kč.	16,701 Kč.	24,727 Kč.	26,528 Kč.	3,421 Zl.	1,437 Dollars	1,341 Dollars
·31	·45	·52	·60	·455	·46	·59	·58	·42	·43	·555	·38	·30
·10	·18	·065	·045	·15	·12	·04	·05	·06	·08	·04	·13	·11 ‡
·125	·14	·19	·15	·115	·12	·095	·10	·12	·11	·17	·17	·18
·06	·05	·045	·05	·06†	·07†	·055	·055	·05	·05	·05	·05	·07
·055	·02	·03	·02	·04	·04	·03	·035	·04	·06	·035	·05	—
·35	·16	·15	·135	·18	·19	·19	·18	·31	·27	·15	·22	·34
·09	·345	·45	·50	·34	·27	·50	·46	·27	·30	·245	·30	·23
·10	·16	·065	·05	·10	·11	·03	·045	·05	·065	·025	·09	·13‡
·12	·26	·255	·21	·13	·16	·15	·14	·14	·135	·28	·24	·19
·05	·04	·025	·03	·07†	·08†	·05	·03	·035	·04	·03	·03	·03
·065	·025	·015	·04	·06	·04	·02	·045	·035	·055	·08	·06	—
·575	·17	·19	·17	·30	·34	·25	·28	·47	·405	·34	·28	·42
0·3	0·8	0·9	0·8	0·7	0·6	0·85	0·8	0·65	0·7	0·45	0·8	0·7
1·0	0·9	1·0	1·1	0·7	0·9	0·7	0·9	0·8	0·8	0·6	0·7	1·2‡
1·0	1·9	1·3	1·4	1·1	1·4	1·6	1·4	1·2	1·2	1·6	1·4	1·1
0·8	0·8	0·6	0·6	1·2 †	1·1†	0·9	0·5	0·7	0·8	0·6	0·5	0·4
1·2	1·25	0·5	1·8	1·5	1·0	0·7	1·3	0·9	0·9	2·2	1·2	—
1·6	1·1	1·2	1·3	1·7	1·8	1·3	1·5	1·5	1·5	2·3	1·3	1·2

* Housekeeping generally.
† Including washing and cleaning materials.
‡ Rent and Furniture merged.

TABLE B. FOOD EXPENDITURE

Expenditures per Equivalent Adult on the Main Food Groups in

	United Kingdom, 1904. Workers.	France, 1904. Workers.	Germany, 1906. Workers.	Belgium, 1908. Workers.	Liverpool, 1929. Workers.	English Towns, 1926. Clerks.	L.S.E., 1932. All Classes.	Belgium, 1928–9. Workers.	Amsterdam. 1918–19. Officials.	Amsterdam. 1923–4. All Classes.	Germany, Workers.	Germany, Salaried Employees.
Number of Budgets	1944	5,605	5,046	1,859	154	90	123	809	82	212	896	542
Average Total Food Expenditure per week	6·0 sh.	6·7 sh.	5·9 sh.	5·4 sh.	8·0 sh.	18·4 sh.	13·8 sh.	78* francs	6·0 florins	6·1 florins	8·3 Marks	10·0 Marks
$\bar{w}$ Cereals . . .	·185	·20	·19	·21	·19	·14	·13	·14	·20	·16	·20	·17
Meat . . .	·31	·33	·35	·29	·30	·34	·26	·27	·20	·19	·29	·28
Dairy Produce .	·19	·19	·23	·22	·25	·26	·30	·32	·30	·30	·27	·28
Vegetables . .	·08	·11	·09	·10	·125	·14	·19	·14	·16	·15	·13	·14
Sugar, etc. . .	·065	·025	·03	·015	·06	·05	·05	·06	·05	·08	·04	·04
Tea, Coffee . .	·07	·05	·05	·05	·07	·045	·05	·06	·03	·04	·03	·04
Miscellaneous .	·10	·10	·07	·12	·01	·025	·02	·01 †	·05	·08	·04	·06
k Cereals . . .	·05	·13	·14	·10	·065	·12	·09	— ·02	·16	·05	·055	·06
Meat . . .	·34	·38	·33	·31	·32	·41	·32	·32	·255	·18	·335	·29
Dairy Produce .	·28	·19	·28	·28	·29	·21	·20	·35	·36	·20	·31	·27
Vegetables . .	·10	·08	·08	·06	·20	·14	·24	·15	·10	·11	·15	·15
Sugar, etc. . .	·06	·01	0	·01	·03	·03	·04	·08	·02	·05	·04	·04
Tea, Coffee . .	·06	·04	·05	·03	·07	·06	·08	·08	·025	·01	·05	·05
Miscellaneous .	·12	·18	·13	·21	·03	·03	·03	·02†	·08	·40	·06	·14
$\bar{\eta}$ Cereals . . .	0·3	0·6	0·7	0·5	0·3	0·9	0·7	— 0·1	0·8	0·3	0·3	0·4
Meat . . .	1·1	1·2	0·9	1·1	1·1	1·2	1·2	1·3	1·3	1·0	1·1	1·1
Dairy Produce .	1·4	1·0	1·2	1·2	1·1	0·8	0·7	1·1	1·2	0·6	1·1	1·0
Vegetables . .	1·2	0·7	0·9	0·6	1·6	1·0	1·3	1·1	0·6	0·7	1·1	1·1
Sugar, etc. . .	0·9	0·4	0	0·9	0·5	0·6	0·8	1·1	0·4	0·7	1·0	1·0
Tea, Coffee . .	0·9	0·9	1·0	0·7	1·0	1·3	1·5	1·3	0·8	0·4	1·4	1·3
Miscellaneous .	1·2	1·9	1·9	1·8	2·6	1·0	1·4	1·6†	1·6	5·3	1·7	2·4

NOTES.—The food groups are composite and have not quite the same content in all budget collections.
Total Food Expenditure excludes alcoholic drinks, tobacco and meals away from home.
Meat includes animal fats such as lard.
Dairy Produce includes margarine with butter, milk and cheese and also includes eggs.
Vegetables include all fresh fruit, often dried fruit and sometimes preserved fruit.
Sugar usually includes honey and sweetmeats.
Tea, Coffee is a group including all non-alcoholic drinks.
Miscellaneous takes salt and spices with a variable group of foods not classified in any of the other groups.

PER EQUIVALENT ADULT

relation to Total Food Expenditure per Equivalent Adult

1927–8.						Sweden, 1923.				Basle.		Poland.	
Officials.	Hamburg, 1925. Workers.	Hamburg and Bremen, 1927–8. Workers.	Copenhagen, 1922. All Classes.	Denmark, 1922. All Classes.	Norway, 1927–8. Workers.	Workers.	Officials.	Middle Class.	Finland, 1920–1. Workers.	1912. All Classes.	1921. All Classes.	1927. Workers.	1929. Workers.
498	65	104	102	141	135	747	445	208	437	78	64	192	84
10·0 Marks	7·8 Marks	9·0 Marks	12·1 Kr.	9·7 Kr.	9·0 Kr.	8·4 Kr.	9·0 Kr.	10·8 Kr.	70·8 Marks	2·5* francs	4·8* francs	9·2 Zl.	9·4 Zl.
·17	·17	·18	·17	·18	·175	·17	·16	·14	·255	·10 §	·10 §	·34	·28
·27	·29	·27	·29	·26	·27	·24	·25	·26	·16	·18	·18	·32	·34
·29	·31‡	·31	·30	·30	·325	·35	·34	·34	·38	·30 ‖	·32 ‖	·13**	·165**
·14	·135	·14	·11	·10	·11	·08	·085	·11	·06	·09	·085	·115	·11
·04	·035	·03	·045	·06	·05	·08	·08	·07	·075	·03	·035	·055	·065
·04	·05	·06	·055	·07	·06	·06	·05	·04	·06	·30	·28	·02	·02
·05	·01	·01	·03	·03	·01	·02	·025	·04	·01			·02	·02
·07	·06	·075	·11	·08	·045	·07	·08	·05	·10	— ·105 §	— ·07 §	·08	·10
·25	·41	·33	·33	·33	·34	·31	·275	·36	·24	·185	·215	·525	·425
·26	·29‡	·365	·27	·27	·38	·37	·375	·25	·42	·21 ‖	·25 ‖	·22**	·275**
·19	·15	·15	·15	·17	·14	·11	·12	·19	·08	·105	·10	·06	·075
·04	0	·005	·015	·03	·02	·055	·055	·05	·07	·02	·02	·055	·065
·04	·08	·07	·065	·065	·05	·05	·055	·03	·075	·58	·48	·03	·03
·14	·01	·005	·06	·06	·02	·035	·04	·07	·015			·03	·03
0·4	0·3	0·4	0·6	0·5	0·3	0·4	0·5	0·3	0·4	— 1·0 §	— 0·7 §	0·2	0·35
0·9	1·4	1·2	1·1	1·3	1·3	1·3	1·1	1·4	1·5	1·0	1·2	1·7	1·25
0·9	0·9‡	1·2	0·9	0·9	1·2	1·1	1·1	0·75	1·1	0·7 ‖	0·8 ‖	1·7**	1·7**
1·4	1·1	1·0	1·4	1·7	1·3	1·3	1·4	1·7	1·3	1·1	1·2	0·5	0·7
1·0	0	0·2	0·3	0·5	0·4	0·7	0·7	0·75	0·9	0·7	0·6	1·0	1·0
1·1	1·7	1·2	1·1	0·9	0·8	0·9	1·0	0·7	1·3	1·9	1·7	1·5	1·3
2·5	0·9	0·6	1·9	2·0	1·7	1·5	1·7	2·0	1·5			1·6	1·3

* Per "quet" of 3·5 equivalent adults.
† Includes wine.
‡ Includes lard.
§ Bread only.
‖ Excludes cheese.
** Excludes margarine.

TABLE C. FOOD EXPENDITURE PER EQUIVALENT ADULT

Expenditures per Equivalent Adult on Certain Food Items in relation to Total Food Expenditure per Equivalent Adult

	$\bar{w}$					k					$\bar{\eta}$				
	Belgium, 1928–9. Workers.	Germany, 1927–8. Workers.	Hamburg and Bremen, 1927–8. Workers.	Sweden, 1923. Workers.	Finland, 1920–1. Workers.	Belgium, 1928–9. Workers.	Germany, 1927–8. Workers.	Hamburg and Bremen, 1927–8. Workers.	Sweden, 1923. Workers.	Finland, 1920–1. Workers.	Belgium, 1928–9. Workers.	Germany, 1927–8. Workers.	Hamburg and Bremen, 1927–8. Workers.	Sweden, 1923. Workers.	Finland, 1920–1. Workers.
Rye Bread	—	·091	·087	·079	·052	—	— ·054	— ·023	·093	·053	—	— 0·6	— 0·25	1·2	1·0
Wheat Bread	—	·039	·028			—	·042	·034			—	1·1	1·2		
Flour	—	·020	—	·080	·160	—	0	—	— ·023	— ·077	—	0	—	— 0·3	— 0·5
Milk	·078	·108	·102	·166	·198	·048	·138	·089	·158	·225	0·6	1·3	0·9	0·95	1·1
Butter	·147	·058	·067	·082	·156	·260	·136	·175	·126	·148	1·8	2·35	2·6	1·5	1·0
Margarine	·018	·040	·050	·031	·011	— ·026	— ·026	— ·025	— ·010	— ·014	— 1·4	— 0·65	— 0·5	— 0·3	— 1·3
Eggs	·059	·044	—	·042	—	·080	·072	—	·063	—	1·35	1·6	—	1·5	—
Potatoes	·055	·048	·050	·030	·034	·013	0	·016	·007	·014	0·2	0	0·3	0·2	0·4
Vegetables	·064	·042	·048	·049	·020	·079	·061	·071	*	*	1·2	1·45	1·5	—	—
Fresh Fruit	·018	·035	·023			·041	*	·030			2·3	—	1·3	—	—
Sugar	·017	·028	·028	·078	·059	·010	·013	·006	·061	·049	0·6	0·5	0·2	0·8	0·8
Coffee	·052	·029	·037	·050	·056	·051	·054	·051	·044	·065	1·0	1·9	1·35	0·9	1·2

* The linear expenditure relation for Fresh Fruit in the German data changes direction at about the centre of the income range, the value of k changing from ·063 to ·090. The same is true of expenditure on vegetables and fruit (together) in Sweden and Finland, the value of k changing from ·075 to ·143 (Sweden) and from ·026 to ·061 (Finland).

ber of equivalent adults as correctly as the data allow in each case. Table C is similar to Table B, but the expenditures are not on food groups but on more specific items of food.

The values of $\bar{w}$ can always be obtained directly from the reports of the budget enquiries and are usually explicitly given in these reports. The values of k, on the other hand, are the result of the fitting of expenditure lines to the data. Where the details of the expenditures of all families are known and where the budgets have been utilised in the following chapter, the value of k has been computed by the usual mathematical method of correlation and regression. In the other cases, it has been necessary to construct graphs, such as have been illustrated in the Diagrams II–VII above, and to read off the values of k from them. Since the data are always more or less imperfect and the positions of the lines somewhat arbitrary, the values of k can only be approximate. They are entered in the tables to two decimal places for convenience but, in general, only the first figure should be considered definite. When the number of budgets is large (500 to 1,000) or when the proportion of expenditure on the particular item is small, it may be possible to use also the second figure. In any case, the significance of the values of k depends on the extent to which the linear law applies to the data and this can only be judged by inspection of the actual graphs. Since it is not possible to reproduce here further examples, we must content ourselves with a few general remarks on the nature of the graphs that have been drawn.

The linear expenditure law suggested by theory can only be expected for a class of families homogeneous as regards tastes and needs and making their purchases on the same market. If a budget collection is to exhibit a definite expenditure relation in practice, therefore, we require that the budgets should be collected as typical of a defined class of families, that they should relate to a town or district in which prices are uniform and that they should provide the information necessary to

eliminate the effect of variable needs. These requirements are very imperfectly realised in many of the actual budget collections handled.

In the first place, budgets are not always carefully separated as regards social class. A large number of collections relate to working-class families only; others relate to salaried employees or to middle-class families. But many of the collections cover either the undefined and heterogeneous class of "Officials" or families of all social classes from workers to professional men. In such cases, tastes and consumption habits vary considerably and cannot be expected to conform to an average. Secondly, many collections of budgets were obtained, not from families in one town, but from towns of varying size and location, while a few collections even included rural or semi-rural areas as well. In addition to variation of tastes between (say) workers in large and small towns, it is now possible that prices vary considerably over the area covered. Thirdly, as we have already noted, it is not always possible to eliminate the effect of the variation of needs as completely as is required.

For these reasons, most budget collections are to be expected to show divergences from the linear law. These divergences are magnified in a number of cases by the fact that only a small number of budgets were collected. Many budget collections contain fewer than 200, and some fewer than 100, families. Considerable sampling errors must, therefore, be present in the distribution of average expenditures.

Taking the range of incomes as a whole, the most irregular distributions of average expenditures were found in the case of the Basle budgets of 1912 and 1921. The number of families is here very small (78 budgets in 1912 and 64 in 1921) and they belong to all social classes. The distribution of the family expenditures of the Belgian workers in 1928–9 also shows considerable divergences. There are only 116 budgets in this case, distributed over all parts of the country. In the case of the food expenditure in the same enquiry, the number

of budgets is much larger and the distribution quite regular. In the Denmark enquiry of 1922, the distribution of food expenditure is better for the Copenhagen families than for the families in the provincial towns taken together. In both groups the number of families is small and all social classes are included. It can be concluded that lack of uniformity in the social class or location of the families decreases the regularity of the expenditure distribution. On the other hand, it is surprising how little effect the smallness of numbers has when the data is fairly perfect in other respects.

When the number of budgets is large, the applicability of the linear expenditure law to a collection including families of very different incomes and social positions can be tested. The U.S.A. collection of 1918 has already been instanced. Other cases are provided by the collections relating to classes of "Officials." Considerable divergences from regularity were found in the distribution of expenditures of the German Officials (1927–8). This class cannot, therefore, be a homogeneous one and, on examination, it was found that it included workers of many social positions, from railwaymen to university professors and highly specialised professional workers. Similar remarks would appear to apply to the official class in Czechoslovakia. On the other hand, as noted above, the officials in Sweden appear to form a class fairly homogeneous as regards tastes.

Where the budgets were both numerous and fairly homogeneous, the graphs were found to exhibit well-defined expenditure relations which were, in the vast majority of cases, of the special linear form tested. Some evidence of curvilinear relations was, however, found in certain cases. The relation of food expenditure to total expenditure appears to be slightly curvilinear, not only in the case of the German workers shown above, but also in the cases of workers in Czechoslovakia (1929) and in Poland (1929). The increase of food expenditure as income increases tends to become less marked in the

higher income groups. It must be remembered, however, that the effect of variation of needs has not been eliminated in these cases. Further, the meat and dairy food groups of expenditure per equivalent adult were also found to be related to total food expenditure in a curvilinear way in some cases. The nature of the relations is seen in Diagram IV above, where, however, it may be due to imperfect elimination of needs. The same kind of curvilinear relation is found more definitely in working-class families in Belgium and Finland. Curvilinear relations, of this slight kind, can only be verified by taking a finer income grouping.

Some interesting graphs were obtained for expenditures on special food items in the four groups of working-class budgets which were sufficiently numerous. Good linear relations were found in most cases together with evidence of a change in direction for some items. In Sweden and Finland, for example, expenditure on bread changes direction upwards, and expenditure on sugar downwards, in the top income groups. The expenditure on vegetables and fruit also changes direction about the middle of the income range, a fact which is supported by the German figures for fruit illustrated above. Working-class families, though homogeneous in respect of tastes for the food groups, are not necessarily homogeneous also for items such as bread, sugar or fruit. Further examination of the expenditures of (say) skilled and unskilled workers on such items would seem to be desirable.

Only very general comparisons should be attempted between the values of $\bar{w}$, k and $\bar{\eta}$ for the different budget collections. The main difficulty in the way of comparison is the fact that the distribution of expenditure on different items is governed, not only by tastes, but also by the relative prices of the items. Budget collections at different dates and in different countries vary in respect of relative prices. Nevertheless, the tables above provide material for interesting, if tentative, comparative conclusions on the distribution of expenditure at different

dates, in different countries and for different social classes.

The clearest comparisons are between the German and Swedish figures for three social classes, the figures being collected at the same time and with a uniform classification. Fairly regular changes are seen as we pass from the German workers to the salaried employees and then to the officials. The proportion, and rate of increase, of expenditure on rent rises, and on food falls, as we pass through the three classes. The proportions spent on various food groups remain constant to a surprising extent but the values of k for (e.g.) meat and dairy produce change regularly.

The main result of a comparison between the British, Belgian and German figures before and after the war is the increased consumption, as measured by $\bar{w}$ and k, of dairy produce and of vegetables. This may be due partly to a change of tastes, but may also be due to a change in relative prices.

The values of k show remarkable general agreement for the different budget collections, particularly in the distribution of food expenditure. The values of k for sugar and coffee are quite small and of the order of $\cdot 05$. The values for cereals vary little about $\cdot 1$. For meat and dairy produce, the values of k are much higher and generally of the order of $\cdot 3$. The same general agreement is noticed between the four budget collections which were examined in respect of the special food expenditures. Negative values of k are found for rye bread, for flour and for margarine. The expenditures on these commodities decrease absolutely as income increases. The value of k for potatoes is small and expenditure almost constant for all incomes. On the other hand, relatively large values of k are found for milk and butter, two items to which large proportions of income are devoted.

The values of income elasticity of demand are also in very good agreement between different budget collections. These values, as we have seen, describe an aspect

of the order of urgency of items and some remarks on the comparison of the different orders for different budgets are given below.

In conclusion, whether comparisons are valid or not, we maintain that a tabulation of the values of $\bar{w}$, k and $\bar{\eta}$, for each budget collection separately, expresses the facts in as clear and concise a manner as is possible.

7. A Graphical Representation of the Distribution of Expenditure

The significance of the values given in the tables above as regards the distribution of expenditure as a whole can be well brought out by a special graphical method of which examples are drawn in Diagrams VIII–XII. The basis of this method is the definition of the order of urgency of the various items in the family budget.

The value of the constant c of the linear expenditure relation governs the position of the line drawn in the previous set of diagrams (as opposed to its gradient) and measures the hypothetical expenditure on the item concerned at zero income. We have seen that, as c decreases from large positive to large negative values, the order of urgency of needs for the items of the budget is established, c being positive for necessaries and negative for luxuries. In the latter case, the expenditure on the item does not begin until some way up the scale of incomes. Instead of c, we could use the expression $(\bar{w} - k)$, which is simply the value of c expressed as a proportion of the average total expenditure.

The graphical method is best appreciated by considering Diagram VIII. This refers to the London School of Economics budgets (1932) and the order of urgency of the items of family expenditure is Housekeeping ($c = £56$), Rent and Rates ($c = £20\frac{1}{2}$), Fuel and Light ($c = £10\frac{1}{2}$), Clothing ($c = £1$) and Other Items ($c = -£88$).

Total expenditure per family is measured on the hori-

DIAGRAM IX.

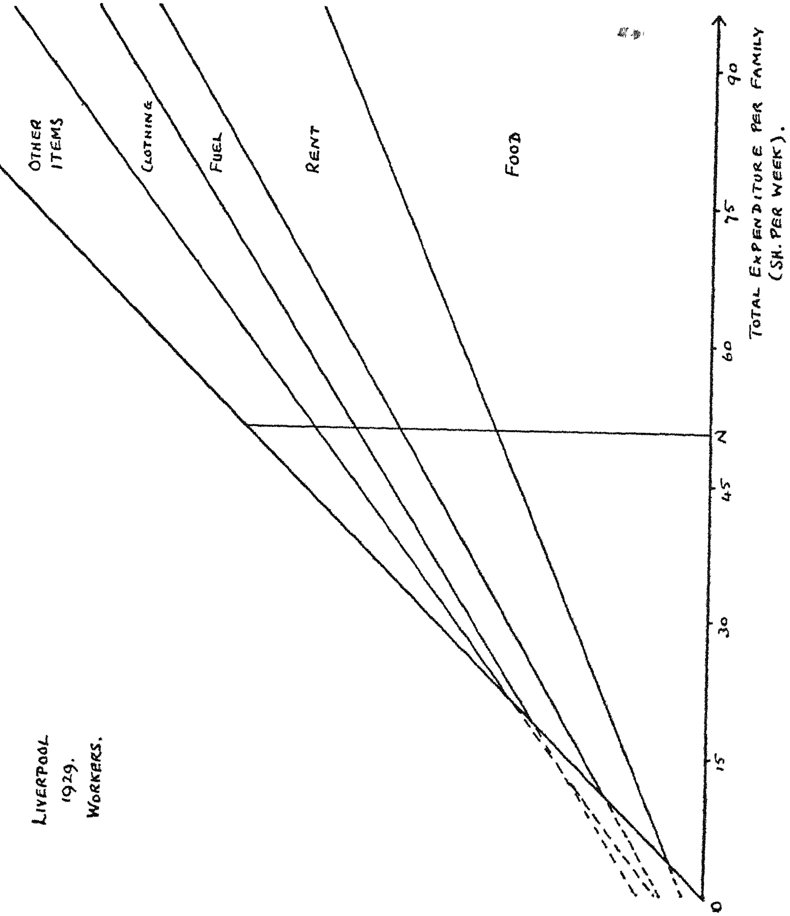
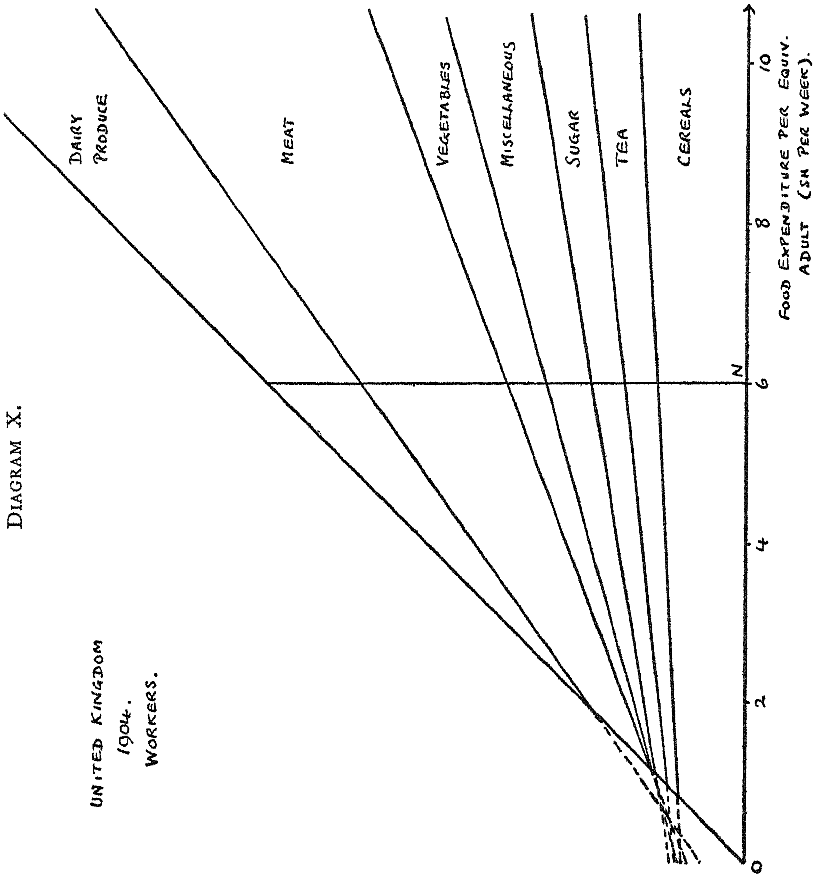

LIVERPOOL
1929.
WORKERS.
OTHER
ITEMS
CLOTHING
FUEL
RENT
FOOD
0
15
30
45
N
60
75
90
TOTAL EXPENDITURE PER FAMILY
(SH. PER WEEK).

DIAGRAM X.


UNITED KINGDOM
1904.
WORKERS.
DAIRY PRODUCE
MEAT
VEGETABLES
MISCELLANEOUS
SUGAR
TEA
CEREALS
N
0
2
4
6
8
10
FOOD EXPENDITURE PER EQUIV. ADULT (SH PER WEEK).

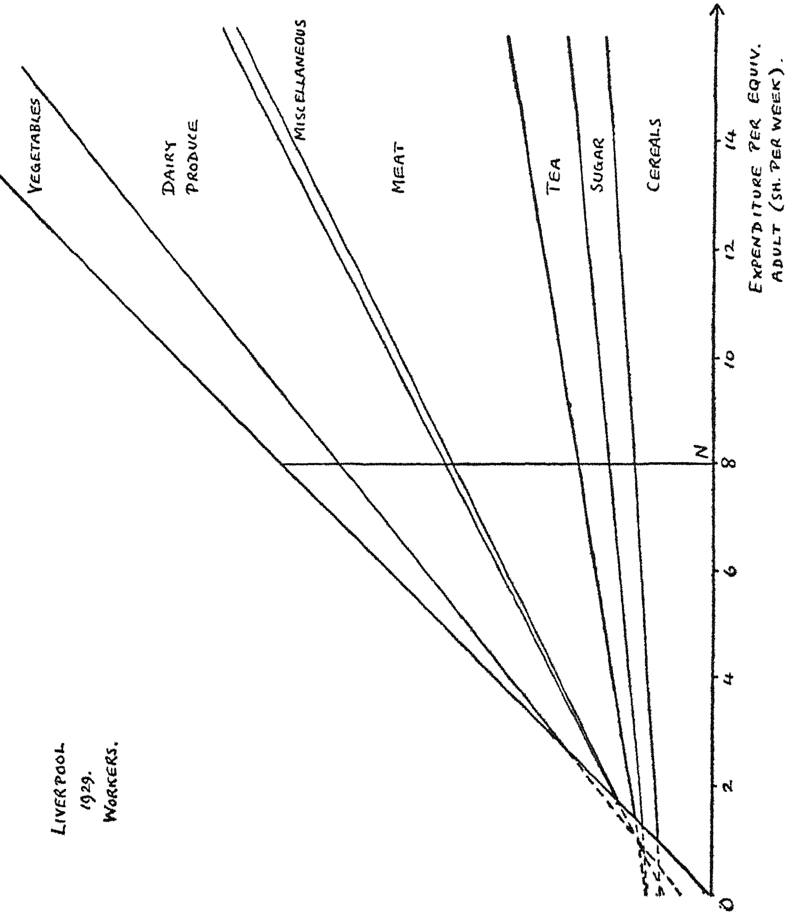
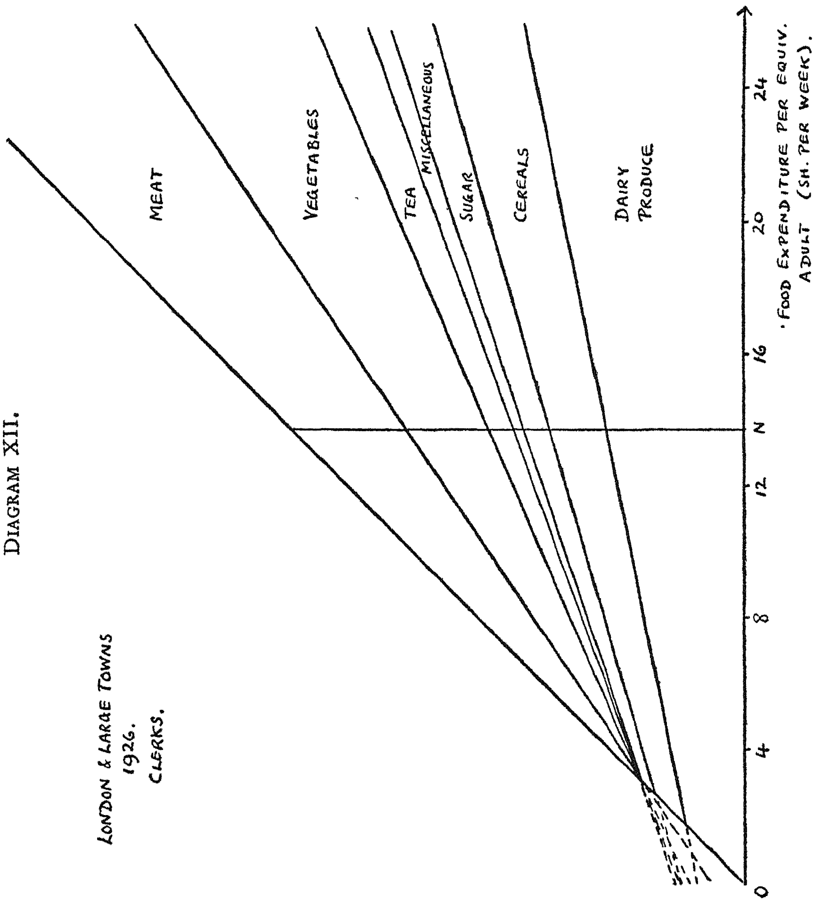

LIVERPOOL
1929.
WORKERS.
VEGETABLES
DAIRY PRODUCE
MISCELLANEOUS
MEAT
TEA
SUGAR
CEREALS
N
0
2
4
6
8
10
12
14
EXPENDITURE PER EQUIV. ADULT (SH. PER WEEK).

DIAGRAM XI.

DIAGRAM XII.


LONDON & LARGE TOWNS
1926.
CLERKS.
MEAT
VEGETABLES
TEA
MISCELLANEOUS
SUGAR
CEREALS
DAIRY PRODUCE
0
4
8
12
N
16
20
24
·FOOD EXPENDITURE PER EQUIV. ADULT (SH. PER WEEK).

zontal axis. Housekeeping is then measured on the vertical axis and represented by a line according to the formula $y = ke + c$, where e is total expenditure. The constants of this equation are found from the complete data by the regression method. If ON is the average total expenditure and NP the average housekeeping expenditure, then the line passes through the point P. A line representing rent *added* to housekeeping is then drawn, a formula of the same kind, as determined from the data by the regression method, being used. If PQ is the average rental expenditure, the second line passes through Q. Other lines are drawn to represent the addition of fuel and light (average expenditure QR) and of clothing (average expenditure RS) to the previous items. Finally the group of other items is added (average expenditure ST) to give the total expenditure line which is necessarily inclined at an angle of 45° to the horizontal.

From the diagram so constructed, the expenditures on the groups of items (in the order of urgency) at any other total expenditure ON_1 are read off as P_1Q_1, Q_1R_1, R_1S_1 and S_1T_1. The actual amounts spent on the various items are thus given by a cross section of the diagram at the appropriate total expenditure level. Unless the expenditure on any group diminishes absolutely as total expenditure rises, the lines diverge as we look to the right. On the other hand, the *proportionate* expenditure on a group diminishes with rising total expenditure in each case where c is positive and rises in each case where c is negative. An increasing proportionate expenditure is only shown when the line concerned intersects the line immediately below it within the diagram. In Diagram VIII, there is a diminution of proportionate expenditure except in the case of expenditure on the other items.

The other four diagrams are drawn in the same way and, with the exception of Diagram X, equations obtained from the data by the method of regression are used in all cases. The values of c, expressed as propor-

tions of the corresponding average total expenditure, *ON*, and written in the order of urgency, are as follows:

DIAGRAM VIII

Housekeeping	·13
Rent	·05
Fuel and light	·02
Clothing	·00
Other items	— ·20

DIAGRAM IX

Food	·05	equal
Rent	·05	equal
Fuel and light	·045	
Clothing	— ·04	
Other items	— ·105	

DIAGRAM X

Cereals	·135
Tea, etc.	·01
Sugar	·005
Miscellaneous	— ·01
Vegetables	— ·02
Meat	— ·03
Dairy produce	— ·09

DIAGRAM XI

Cereals	·125
Sugar	·03
Tea, etc.	0
Meat	— ·02
Miscellaneous	— ·02
Dairy produce	— ·04
Vegetables	— ·075

DIAGRAM XII

Dairy produce	·10
Cereals	·04
Sugar	·01
Miscellaneous	— ·01
Tea, etc.	— ·03
Vegetables	— ·05
Meat	— ·06

These five diagrams enable us to compare the order of urgency of items in the British groups of family budgets we have examined. This comparison can be extended to other budget collections and a number of interesting conclusions can be drawn.

The order of urgency of the main groups of family expenditure varies little in the budget collections considered. Food is the expenditure of first necessity, followed by rent and then by fuel and light. Furniture and clothing, the next items in the order, are definite luxuries, and finally comes the miscellaneous group of expenditure. There are no notable exceptions to this order. In the case of the Liverpool workers, rent is an expenditure of approximately the same urgency as food and fuel is only a little less urgent. All the other budget collections show food as definitely more urgent

than rent. There is some indication, in many cases, that clothing is an expenditure of less necessity for the working-class families.

The main groups of food expenditure vary considerably, in respect of the order of urgency, at different dates, in different countries and particularly for different social classes. It is possible, however, to discern a fairly established order for the groups of working-class families examined. The order puts the cereal group as the expenditure of the first necessity, followed by sugar and coffee or tea with dairy produce and meat as very definite luxuries. The position of the vegetable group in the scale varies a good deal in different countries. For the Liverpool workers, vegetables are a greater luxury than even meat or dairy produce, as is shown by Diagram XI. The vegetable group is also a luxury for German and Scandinavian workers. On the other hand, it is a group of marked necessity in the cases of the Belgian and Polish workers, and also in the case of the official class in Amsterdam (1918–19). Comparing the Liverpool working-class budgets of 1929 with the more general British budgets for workers in 1904, it appears that meat expenditure has become more, and vegetable expenditure less, necessary since 1904. This is shown by Diagrams X and XI. The reverse appears to be the case for German workers over the same period.

In three countries, the United Kingdom, Germany and Sweden, it is possible to compare the order of urgency of food items for workers with that for middle-class families. Significant differences are found and the position of the dairy group should be particularly noticed. Dairy produce is a definite luxury for workers but an expenditure of necessity for the middle-classes. This is true of the British families (see Diagrams XI and XII) as well as of the German and Swedish families. There is also some evidence that expenditure on beverages (mainly tea for British families and coffee for continental families) is more urgent for working-class than for middle-class families.

In comparing the orders of urgency established in different budget collections, it should be remembered that the order depends on the relative prices of the items as well as on the tastes of the families. It can be noticed, in illustration of this, that dairy produce appears higher in the scale of urgency in countries, such as Denmark and Switzerland, which specialise in dairy production.

8. The Budget Collections Utilised

The collections of family budgets of which use has been made in the numerical illustrations of the present chapter are listed below.[1] The expenditures relate to a whole year in all cases except those specially noted in the list.

1. *United Kingdom*, 1904. 1,944 workers.
   France, 1904. 5,605 workers.
   Germany, 1906. 5,046 workers.
   Belgium, 1908. 1,859 workers.

These enquiries were conducted by the British Board of Trade on a nearly uniform basis. They were designed to afford comparison of prices and standard of living amongst workers in the different countries. The expenditures relate to one week in all cases.

The first collection provides the foundation of the weighting used in the British Cost of Living index-number. These weights, as modified to apply to 1914, are recapitulated in Table A.

2. *Liverpool*, 1929. 154 workers.
   The Social Survey of Merseyside (ed. D. Caradog Jones), Vol. I (Liverpool, 1934).

[1] We are greatly indebted to Dr. Hans Staehle for his table of family budget enquiries published in *Econometrica*, January, 1935. This table contain invaluable information on the nature and scope of budget enquiries made in European countries since the war and we have made much use of it in selecting the budget collections suitable for our purpose.

Our selection of budget data excludes, almost entirely, the numerous enquiries made in the U.S.A. Information on the nature of these enquiries is to be obtained from the bibliographies issued by the U.S. Department of Agriculture.

These working-class budgets were collected in connection with a general survey of social conditions in Liverpool and neighbourhood. The expenditures relate to one week only, and different families submitted budgets for different weeks at various times in the year. The budgets relate mainly, but not exclusively, to 1929.

By courtesy of Mr. Caradog Jones, we have been able to use the original returns in full detail. Of the complete budget collection, however, only those cases were retained where the head of the family was in full-time work in the week in question.

3. *English towns*, 1926. 112 clerks in London and other large towns; 122 clerks in smaller towns.

Caradog Jones, "Cost of Living of a Sample of Middle-class Families," *Journal of the Royal Statistical Society*, 1928.

Actual or estimated annual figures for the main groups of family expenditure are given. The food expenditures relate to the month of February, 1926, and are given in the cases of only 90 families in London and large towns and 104 families in smaller towns. Mr. Caradog Jones has also lent us the original returns obtained in this collection of budgets.

4. *London School of Economics*, 1932. 123 families of all classes.

These budgets were obtained in connection with the enquiry into *Family Life* organised by Sir William Beveridge in co-operation with the B.B.C. They were submitted voluntarily in response to broadcast appeals. The results have not hitherto been published and are, therefore, set out at some length in this and the following chapter.

5. *Belgium*, 1928–9. 116 workers (family expenditure); 809 workers (food expenditure).

Armand Julin, "Résultats principaux d'une enquête

sur les budgets d'ouvriers et d'employés en Belgique (1928–9)," *Bulletin de l'Institut International de Statistique* (XXII[e] Session, London, 1934).

The food expenditures relate to four fortnights in the year April, 1928–March, 1929, one fortnight at the beginning of each quarter. The family expenditures relate to the whole year but were obtained from a smaller number of families. Budgets were also collected from families of salaried employees but no use of them has been made here.

6. *Amsterdam*, 1918–19. 82 officials.

Statistische Mededeelingen van het Bureau van Statistiek der Gemeente Amsterdam, No. 73, *De Uitgaven van* 114 *Ambtenaars- en Arbeidergezinnen* (Amsterdam, 1924).

The budgets relate to September, 1918–September, 1919. Budgets from 32 working-class families were also collected in the same enquiry but they have not been used here.

7. *Amsterdam*, 1923–4. 212 families of all classes.

Statistische Mededeelingen van het Bureau van Statistiek der Gemeente Amsterdam, No. 80, *Huishoudrekeningen van* 212 *Gezinnen uit verschillende Kringen der Bevolking* (Amsterdam, 1927).

The budgets relate to October, 1923–September, 1924. Miss Ouwerkerk has kindly allowed us to use her computations on food expenditure from the published detailed budgets in this collection.

8. *Germany*, 1927–8. 896 workers; 542 salaried employees; 498 officials.

Einzelschriften zur Statistik des Deutschen Reiche, Nr. 22, *Die Lebenshaltung von* 2000 *Arbeiter-, Angestellten- und Beamtenhaushaltung* (Berlin, 1932).

The budgets relate to March, 1927–February, 1928, and were collected in a large number of towns of all sizes.

The workers (*Arbeiter*), salaried employees (*Angestellter*) and officials (*Beamten*) are treated separately.

9. *Hamburg*, 1925. 65 workers.
Statistische Mitteilungen über den hamburgischen Staat, Nr. 20, *Die Lebenshaltung minderbemittelter Familien in Hamburg im Jahre* 1925 (Hamburg, 1926).
Budgets were also collected from 13 families of the salaried employees class but these have not been utilised here.

10. *Hamburg and Bremen*, 1927–8. 104 workers.
These budgets relate to the working-class families of Hamburg (including Altona) and Bremen comprised in the larger German enquiry of 1927–8 listed above (No. 8).

11. *Denmark*, 1922. 102 families of all classes in Copenhagen; 141 families of all classes in provincial towns.
Denmarks Statistik, Statistiske Meddelelser, 4 Raekke, 69 Bind, 5 Haefte, *Husholdningsregnskaber for* 1922 (Copenhagen, 1925).
The families are largely of the official class but smaller proportions of working-class and middle-class families are also included. The budgets relating to Copenhagen and those relating to the provincial towns are considered separately.

12. *Oslo*, 1912–13. 85 families of all classes.
Husholdningsregnskaper ført av endel mendre bemidlede familier i Kristiania etc. i aaret 1912–13 (Oslo, 1915).
The complete collection includes budgets from families in Norwegian towns other than Oslo but only the Oslo budgets have been used here.

13. *Oslo and Bergen*, 1918–19. 75 families of all classes.
Norges Offisielle Statistik, VII, 13, *Husholdningsregnskap* Sept. 1918–Sept. 1919 (Oslo, 1921).

This enquiry again covers families in a number of towns in Norway but only the budgets collected in Oslo and Bergen have been used here.

14. *Norway*, 1927–8. 135 workers.

Norges Offisielle Statistik, VIII, 103, *Husholdnings-regnskap* 1927–1928 (Oslo, 1929).

The budgets relate to September, 1927–August, 1928, and to working-class families in Oslo and four other large towns. Budgets were also collected from 31 families of the official class but these have not been included here.

15. *Sweden*, 1923. 747 workers; 445 officials; 208 middle-class families.

Sveriges Officiella Statistik, Sociastatistik; *Levnad-kostnaderna i städer och industriorter omkring år* 1923 (Stockholm, 1929).

The budgets were collected in a large number of towns and other industrial centres in all parts of Sweden. The three classes are treated separately.

16. *Finland*, 1920–1. 437 workers.

Finlands Officiella Statistik, XXXII; Sociala Special-undersökningar V: *Levnadkostnaderna under Bokförings-perioden*, 1920–21 (Helsingfors, 1925).

This enquiry covers working-class families from industrial and rural areas in all parts of Finland. The complete collection also includes 117 budgets from families of the officials class but these have not been used here.

17. *Basle*, 1912 *and* 1921. 78 families of all classes (1912); 64 families of all classes (1921).

Mitteilungen des Statistischen Amtes des Kantons Basel-Stadt, Nr. 45, *Haushaltsrechnungen von Basler Familien aus den Jahren* 1912, 1919–1923 (Basel, 1925).

Budget enquiries were conducted in each year from 1919 to 1923. The year 1921 has been selected to represent this post-war period.

18. *Czechoslovakia*, 1927 *and* 1929. 79 workers (1927); 262 workers (1929); 122 officials (1927); 291 officials (1929).

Zprávy státního úřadu statistického Republiky Čisko-slovenski, Jg. XII (1931), No. 225–9, and Jg. XIV (1933), No. 138–43.

These budget enquiries have been conducted over a long period of years since the war and the years 1927 and 1929 have been taken to represent the more recent of the enquiries. The expenditures relate to a year ending between July, 1927, and June, 1928, and between July, 1929, and June, 1930, respectively. The working-class families and the officials are considered separately in each year.

19. *Poland*, 1927 *and* 1929. 192 workers (1927); 84 workers (1929).

Budżety Rodzin Robotniczych Statystyka Polski, XL, 1 and 2 (Warsaw, 1930 and 1933).

The budgets relate to working-class families in Warsaw and other industrial regions in Poland. A similar enquiry was conducted in 1928 but no use has here been made of the results.

20. *U.S.A.*, 1918. 12,096 families of all classes.

Bulletin of the United States Bureau of Labour Statistics, No. 357, *Cost of Living in the United States* (1924).

The budgets were collected from all parts of the United States and cover working-class and other families.

21. *U.S.A.*, 1928–9. 269 farmers.

The N.Y. State College of Home Economics at Cornell University has made detailed studies of the household income and expenditure of farmers in Livingstone and Tompkins counties in N.Y. State. Professor Helen Canon has kindly sent us the results in detail and allowed us to use her computations.

In the budget collections numbered 2, 3, 4, 10, and 21 in the list above, the details of expenditure for each family have been available and further use of these details is made in the following chapter. In these cases, as also in the cases of the collections 7 and 9, the expenditure relations of the present chapter are obtained from the full details and not from averages. The published reports of the collections 6, 8, 11, 12, 13 and 19 contain full details of each family's expenditure, but only the averaged tables have been used in the present study.

The food expenditures considered in the present chapter are all reduced to an equivalent adult basis. This elimination of family needs has been done according to one or other of the well-known scales of equivalence. In certain cases, however, the reduction has only been possible on a very rough basis. In the Board of Trade pre-war budget enquiries, family expenditures are given together with the average number of children in each income group. Each child has been taken as half an adult and it has been assumed that two adults are present in each family. In the Hamburg budgets of 1925, each child up to the age of 11 has been counted as half an adult and all other persons reckoned as adults.

Where the published results in average form have been used, it has been necessary, of course, to accept the method of reduction to equivalent adults adopted in the original compilation. In some cases, instead of giving the average expenditure per equivalent adult over the families included in a group, the reports give the total expenditure of all families divided by the total number of equivalent adults in the families. The error involved here is not likely to be great.

In a number of budget collections, the published averaged expenditures are grouped in an inadequate or inconvenient way from the point of view of the present study. For example, in the collections numbered 12, 13 and 18 above, the families are grouped only in grades of total income or expenditure. In these cases, it has not

been possible to use the tables of food expenditure per equivalent adult. On the other hand, in the collections numbered 6, 14 and 15, the families are grouped only in grades of total expenditure per equivalent adult and it has not been possible to consider family expenditure where no reduction on account of needs is made. In other cases, alternative groupings according to total family expenditure and to total expenditure per equivalent adult are usually given.

CHAPTER II

VARIATION OF EXPENDITURE

1. The Problem of Varying Tastes

We have found in Chapter I that average expenditures on the various items of the family budget can be related to average income by simple formulæ applying over a wide field and that these expenditures can be readily classified according to a scale of urgency. Where the budget collections are presented only in average form, we cannot proceed further with their analysis. But we have several collections giving full details of the expenditures of each family and we have used some of these to study the extent and nature of the variations of expenditure about the average.

We first observe, for each budget collection as a whole, the dispersion of the incomes about their general average, i.e. we obtain the frequency distribution of incomes. We deal similarly with expenditures on food, rent, clothing and other items in whatever sub-divisions seem useful. These income and expenditure distributions are markedly unsymmetrical in nearly every case, but it is possible to represent them by a continuous algebraic formula. We next test the effect of expressing income or expenditure per equivalent adult, allowance being thereby made for differences in the age and sex composition of the families. We observe whether the asymmetry of the original distributions is removed in the revised distributions. It is found, in general, that symmetry has not been attained by this process, though it has been improved.

The next stage is to take income (or total expenditure)

as the independent variable and to correlate expenditures on particular items, or groups of items, with income. We can either keep the family as the unit or take account of the number of equivalent persons in the family. In either case, we obtain formulæ of the kind already discussed in Chapter I. Instead of testing the validity of the formulæ by means of the averages in different income groups, we now apply them to each budget separately. We have, for each family, a formula of the kind

$$y_t = ke_t + c + v_t$$

where e_t is the total expenditure (or the expenditure on food as a whole), y_t is the expenditure on a particular item and v_t is the residual expenditure which completes the equality for the tth family. If the effect of variation of needs is being eliminated at the same time, then we either include an additional term dependent on the size of the family (as in the case of rent) or divide expenditures by the number of equivalent adults in the family (as in the case of food expenditures).

The constants k and c are determined, not by the graphic methods of Chapter I, but by fitting a line to the observations by the method of least squares. The expression $(ke_t + c)$ measures the normal or average expenditure on the item of a family whose total expenditure is e_t. The residual v_t measures the variation of the tth family from the norm in respect of the expenditure considered.

In other words, we have eliminated the effect of the amount of income on family expenditure and also, if necessary, the effect of the size of the family. Thus we tend to have now the same needs in relation to the same income and the residual is due to individual variation of tastes or habits.

The most obvious way of examining the variation of the residual expenditure v_t is to compare it with that shown by the normal curve of error. This curve gives the result of a number of fortuitous or unrelated

variations such as we may expect from individual tastes displayed in a uniform field. If the variation of residuals is adequately described by the normal curve of error, we have some evidence, at least, that we have successfully eliminated disturbing factors, such as the amount of income and the extent of needs, and have a uniform field in which there is free play of individuality.

Without further discussion of the implications of the results, we may state that the normal curve of error does, in fact, serve to describe the variations of expenditure in a large number of cases. This will be shown in the sequel.

The meanings of k and c, and of the measures derived from them, have been explained in Chapter I in their application to averages. These meanings can now be taken as applying to the norm. We still need a measurement of the residual variation in addition to the statement that its distribution is "normal." Such a measure is found in its standard deviation which, together with the central value, determines the normal curve in complete detail. The average of the residuals v_t is zero in each case, since the value of the constant c is determined to ensure this. We have then only to state, for each class of expenditure, the standard deviation, say σ_v, of the residual variations. If σ_v is zero, there is no variation and each family spends according to the norm, i.e. y_t is determined when e_t is given. The larger is the value of σ_v, the larger part does individual taste play in the determination of expenditure. We get a useful measure, independent of the units used in expressing expenditures, by taking σ_v as a percentage of the average expenditure of the whole group of families on the item concerned, i.e. by writing what is called the co-efficient of variation. As a result, the expenditure of a group of families on a particular item (such as meat) can be fully described by the constants k and c, together with the derived values of $\bar{w}$ and $\bar{\eta}$, and by V the co-efficient of variation of individual expenditures.

Before proceeding, we may summarise the significance of the standard deviation σ_v from the mathematical side. Write σ_e and σ_y for the standard deviations of e_t and y_t over all families and r for the co-efficient of correlation between them. If $\bar{y}$ and $\bar{x}$ are the average values, then the line fitted to the observations by least squares has constants given by

$$k = r\frac{\sigma_y}{\sigma_e} \qquad \text{and} \qquad c = \bar{y} - r\frac{\sigma_y \bar{x}}{\sigma_e}$$

It then follows, quite easily, that

$$\sigma_v = \sigma_y \sqrt{1 - r^2}$$

The standard deviation of residuals is small when r approaches unity, i.e. when the correlation between expenditure and income approaches perfection, or when σ_y is small, i.e. when the variation of the original expenditures is small.

Hence, when we remove the normal expenditure $(ke_t + c)$ from the original expenditure of each family, the residual variation of expenditure is necessarily smaller than the original variation. The only exception is when $r = 0$, in which case our relation does not help us to establish any new explanation of behaviour.

We have, therefore, to observe the residual expenditures v_t in detail and to see, in the first place, whether the original asymmetry in the distribution of expenditures has been remedied and, in the second place, whether the distribution of the residuals agrees with the normal distribution as nearly as the hazard of sampling leads us to expect. If these conditions are satisfied, we may regard the description of the expenditure of the group as complete. With some hesitation, we may hope that we have analysed all that can be attributed to definite causation and isolated the immeasurable individual element. If the conditions are not satisfied, then there are still some factors of a general character whose effect we have not isolated and neutralised.

2. Frequency Distributions of Income and Expenditures

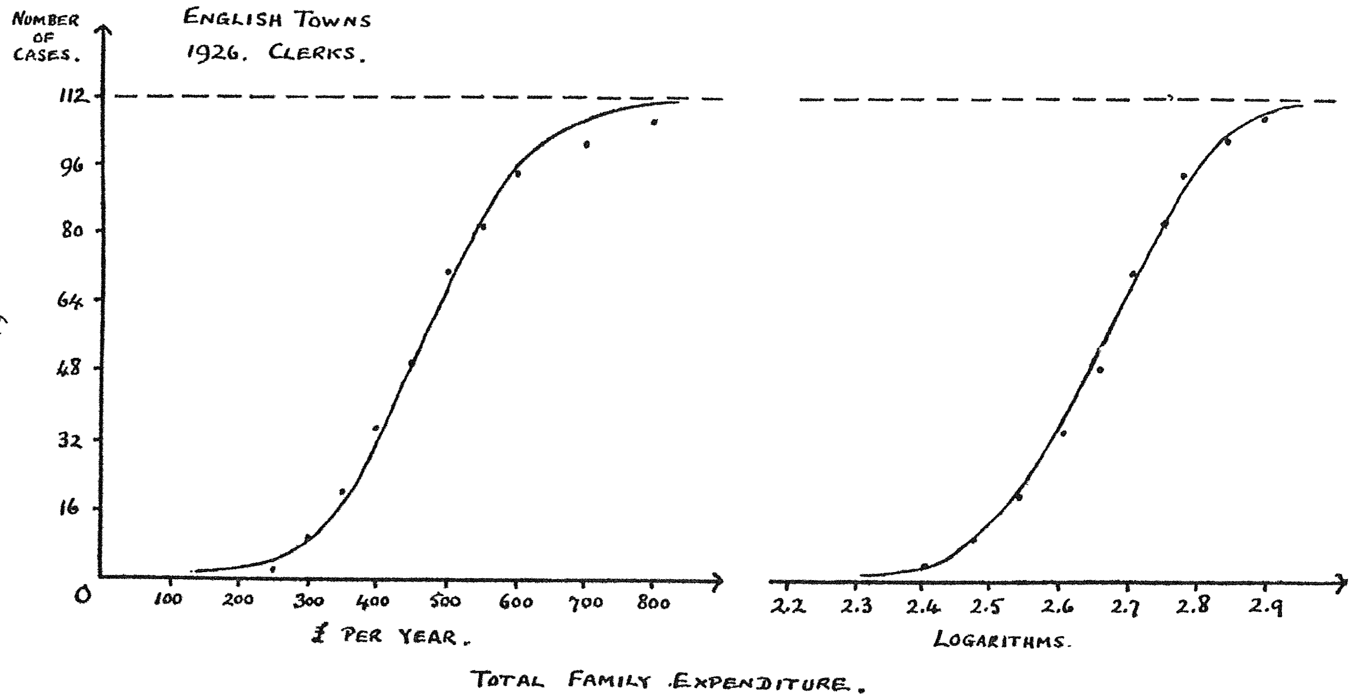
We have examined the distribution about the general average of a large number of characteristics of our detailed budgets—incomes and total expenditures; expenditures on the main groups of family expenditure (food, rent, clothing, etc.); expenditures per equivalent adult on food as a whole and on various food items and groups of items. The great majority of these distributions have a similar form. There is a concentration a little below the average, very few cases far below and a long spread of cases above the average. The typical distribution is illustrated by Diagram XIII, where a distribution of total expenditures is plotted, for convenience, in cumulative form. The height of each plotted point above the horizontal measures the number of families with a total expenditure *not more than* the amount represented on the horizontal axis.

The asymmetry of these distributions is to be expected since there is a natural avoidance of low expenditures below the minimum of subsistence. It is not possible, to name an extreme case, to have negative expenditures, but there is no assignable upper limit to expenditure.

In many cases, the asymmetry can be removed by the simple device of plotting the number of families against expenditures, not on the "natural" scale, but on a scale of logarithms. This is shown in Diagram XIII, where the horizontal scale of the second graph takes the logarithms of total expenditure. The resulting distribution is more regular in the second graph than in the first. The smooth curve drawn in the first graph is the appropriate normal curve of error and the fit is seen to be bad, particularly at low and high levels of total expenditure. On the other hand the same curve fits the distribution of points in the second graph as well as can be expected with so small a number of families. The extreme points now lie quite close to the curve.

The curve traced in the second graph is, in fact, the modification of the normal curve of error derived by using

DIAGRAM XIII.



the logarithm of the independent variable (total expenditure in this case) instead of the variable itself.[1] It is maintained that this modified normal curve is an adequate fit to the asymmetrical distributions of expenditures in a group of families.

Whatever the shape of the distribution, the average and standard deviation of expenditures, together with the resulting co-efficient of variation, can be computed and have their importance. The averages, which have been sufficiently discussed in Chapter I, are used for brief descriptions of the groups of families and especially in the computations of cost of living index-numbers. The standard deviations or the co-efficients of variation give a first measure of the dispersions of the observations about the general averages. Many of the co-efficients of variation we have found are given in Table D. It is noticeable that they are of considerable magnitude. The smallest is 19 per cent for a group of middle-class families in small English towns and relates to total expenditure per equivalent adult. The largest is 117 per cent for family expenditure on clothing in the case of working-class families in Liverpool. This high value is not surprising since the budgets relate to expenditure in one week only and the purchase of clothing in that week involves a fortuitous element.

As might be expected, the variation in expenditures reduced to the basis of equivalent adults is less than that

[1] The equation of the normal curve can be written

$$y = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{(x-a)^2}{\sigma^2}}$$

where a is the average and σ the standard deviation of the observations. Writing z for $\log x$, the modified normal curve has equation

$$y = \frac{b}{\sqrt{\pi}} e^{-(bz-c)^2}$$

where b and c are constants to be determined from the data. The constants are best obtained by a graphic method described by M. R. Gibrat in *Les Inegalités Économiques* (1931). M. Gibrat names the logarithmic normal equation *La loi de l'effet proportionel.* The fundamental idea is that equal importance is given to the *proportional* variations in the variable x, instead of to *absolute* variations as in the normal curve of error itself.

TABLE D. CO-EFFICIENTS OF VARIATION OF EXPENDITURE

	Liverpool, 1929. Workers.	English Towns, 1926. Clerks.		L.S.E., 1932. All Classes.
		London and Large Towns.	Small Towns.	
Family Expenditure:				
Total Expenditure	47	31	36	60
Housekeeping	—	—	—	29
Rent	38	31	33	—
Clothing	117	—	45	53
Fuel and Light	29	—	—	35
Expenditure per equivalent adult. Food Groups:				
Total Expenditure	43	24	19	43
Total Food Expenditure	37	24	24	26
Cereals	33	40	—	40
Meat	48	35	35	43
Dairy Produce	51	25	23	28
Vegetables	72	35	—	44
Sugar, etc.	52	33	—	49
Tea, Coffee	60	41	—	59
Expenditure per equivalent adult. Food items:				
Bread	38	32	38	—
Bacon	90	50	48	—
Sugar	35	33	45	—
Tea	48	50	65	—

	Hamburg and Bremen, 1927–8. Workers.
Expenditure per equivalent adult. Food items:	
Total Food Expenditure	19·3
Rye Bread	30·4
Wheat Bread	59·4
Milk	38·2
Butter	78·3
Margarine	56·2
Potatoes	30·0
Vegetables	46·7
Fresh Fruit	64·2
Sugar	37·0
Coffee	63·9

For U.S.A. 1928–9 (Farmers), the co-efficient of variation of Total Expenditure is 50 per cent, and of Food Expenditure per equivalent adult, 45 per cent.

of expenditures per family. The variation of food expenditure is less than that of total expenditure. When we look at the separate food groups, we find an increase in variation and considerable irregularity from group to group and from one set of budgets to another.

It might be thought, at first sight, that such considerable variation renders it impossible to make any useful generalisation, and that cost of living index-numbers are consequently of very limited application, if indeed they are not actually misleading. It may be remarked, in passing, that the variations found in the sets of budgets under discussion are not due to the numbers of budgets being so small; the variations are, in fact, mainly independent of the number of budgets taken. This criticism of the index-numbers would have some validity if we were considering how closely an observed change in a general index-number corresponded to the change in the purchasing power of the income of an individual family. But even that would depend on the effect of differing "weights" on the particular price movements in question, and, in ordinary circumstances, this effect is trifling.

When, however, we are considering the accuracy of averages, and the accuracy of a measurement designed, to quote the Ministry of Labour's standing rubric, to show the average increase in the cost of maintaining unchanged the pre-war standard of living of working-class families, our ground is more secure, especially if we insert the word average before working-class. It is known that the accuracy of an average increases in proportion to the square root of the number of cases on which it is based, provided that the cases are selected at random. This must be interpreted, however, as applying only to the class and locality from which the budgets were obtained.

Thus with only 64 budgets, the minimum number treated here, the co-efficient of variation of an average is one-eighth of that of the group. If we take 40 per cent as a typical co-efficient of variation of the expenditures in a group, the co-efficient of the average should be written

as 5 per cent. Index-numbers are based on a minimum of, perhaps, 2,000 budgets and, with such a number, the co-efficient of variation of an average becomes less than 1 per cent.[1]

3. Elimination of Variation due to Income

The want of symmetry in the distributions of expenditures discussed above is largely due to the fundamental asymmetry of the income distribution. In Chapter I, many cases were given which showed that the average expenditure on an item in a particular income grade was connected with that income approximately by a linear relation. Eliminating the effect of income when we have detailed information for each budget, we write, for the tth family,

$$y_t = ke_t + c + v_t$$

where y_t and e_t are the expenditures of the family on the particular item and on all items respectively and where the constants k and c are determined by the least squares method from the separate budgets. The quantity v_t is due to individual variation of tastes, being the difference between the actual expenditure and that given by the average linear relation between the special expenditure and income. We proceed to study the behaviour of v_t.

There are two things to consider; the co-efficient of variation of v_t and the distribution of v_t about its average (which is necessarily zero).

[1] In computing an index-number, the co-efficient of variation of average expenditures is reduced if the correlation between the expenditure on a particular item and that on all items is considerable. Write V_e and V_y for the co-efficients of variation of total expenditure and of expenditure on any item, and r for the co-efficient of correlation between the special and general expenditures. As in Chap. I, ϖ denotes the ratio of the special to the general expenditure when the budgets are averaged. Then V_w, the co-efficient of variation of ϖ, is approximately

$$\sqrt{V_e^2 + V_y^2 - 2rV_eV_y} \Big/ \sqrt{n}$$

For example, in the Liverpool working-class budgets, V_e is 47 per cent, V_y is 38 per cent for rent and r is ·65.

Then $\qquad V_w = 37/\sqrt{n} = 3{\cdot}7$ per cent $\qquad$ if $n = 100$.

TABLE E. CO-EFFICIENTS OF VARIATION OF RESIDUAL EXPENDITURE

Family Expenditure.	In relation to Total Expenditure.		In relation to Total Expenditure and Size of Family.	
	Rent.	Food.	Rent.	Food.
Liverpool, 1929. Workers	30	—	30	—
English towns, 1926. Clerks	26	—	—	—
L.S.E., 1932. All classes	—	25	—	20
U.S.A., 1928–9. Farmers	—	10	—	9

Expenditure per Equivalent Adult, in relation to Total Expenditure or to Total Food Expenditure.	Liverpool, 1929. Workers.	English Towns, 1926. Clerks.	L.S.E., 1932. All Classes.
Total Food Expenditure	23	22	22
Food Groups:			
Cereals	31	40	—
Meat	37	34	—
Dairy Produce	32	24	—
Vegetables	50	33	29
Sugar, etc.	48	33	44
Tea, Coffee	56	40	46
Food Items:			
Bread	36	—	—
Bacon	82	—	—
Tea	42	—	—

Expenditure per Equivalent Adult, in relation to Total Food Expenditure.	Hamburg and Bremen, 1927–8. Workers.
Food Items:	
Rye Bread	30·1
Wheat Bread	56·6
Milk	34·2
Butter	74·5
Margarine	55·5
Potatoes	29·3
Vegetables	37·6
Fresh Fruit	59·2
Sugar	37·0
Coffee	58·2

For L.S.E., 1932, in relation to total expenditure, the coefficient of variation of rent is 39 per cent, of clothing 35 per cent, and of fuel 29 per cent (all expenditures per equivalent adult).

We have stated that

$$\sigma_v = \sigma_y \sqrt{1 - r^2}$$

and for particular values of r it follows that

r	0	·1	·2	·3	·4	·5	·6	·7	·8	·9	1
$\frac{\sigma_v}{\sigma_y}$	1	·995	·98	·95	·92	·87	·80	·71	·60	·44	0

Hence, unless r is quite large, the variation of v_t is not much less than that of y_t. It must be taken, in general, that there is little difference between the co-efficients of variation of the residual and the original expenditures. Table E exhibits the values of the co-efficients of variation of residual expenditure in some of the cases we have examined.

Passing to the actual distribution of v_t, the first consideration is that of symmetry. As a preliminary test, it can be determined whether there is approximate equality between the number of positive values of v_t and the number of negative values over the whole set of families considered. Equality is expected since the average of the v_t's is necessarily zero.[1] A more complete test of symmetry can only be made by representing the distributions graphically. In the tables of the following section, the number of positive and of negative values of the residual expenditures are given in comparison with the numbers of values of the original expenditures above and below the average expenditure. From this, an idea of the improvement in symmetry can be derived.

If approximate symmetry is found, we can proceed to compare the distribution of residual expenditures with that given by the normal curve of error. If there is no common cause of variation in expenditure, other than

[1] If there are n budgets, we may write the number of positive (or of negative) values to be expected as

$$\tfrac{1}{2}n \pm \tfrac{1}{2}\sqrt{n}$$

where the second term indicates the standard deviation. So, if n is 150, we have 75 ± 6. If we have a random selection from a symmetrical group, the chance is 2 to 1 that the positive entries will number between 69 and 81.

those due to income and needs, both of which have now been eliminated, then the values of v_t are the result of sporadic preferences. If there is no accounting for tastes, we can nevertheless measure their play by their results in distribution, and complete sporadicity leads quite commonly to the normal distribution. A more detailed consideration of the effect of the variation of tastes on the distribution of expenditures in actual budget collections, and an account of the relation of the observed results to the theoretical postulates, is given at a later stage.[1]

To facilitate the comparison between the actual distribution of the residual expenditures and the normal distribution, a uniform system of tabulation has been adopted in the preparation of the tables given below. The standard deviation of the original or residual expenditures was first calculated. The data were then grouped above and below the average. Finally, the grouping was according as the expenditures fall within intervals indicated by successive third parts of the standard deviation away from the average. This grouping was carried up to a deviation from the average of twice the standard deviation; two extra groups (above and below the average) were then added for expenditures between two and three times the standard deviation and for expenditures greater than three times the standard deviation away from the average. The corresponding entries that would occur in the normal distribution, with the same total number of cases, are given for reference in each table. The choice of these particular intervals has been directed, of course, only by convenience.

When the numbers of cases in question are so small, it is difficult to judge precisely whether the correspondence with the normal curve is as close as would be expected in a random sample. To aid judgement in this respect, the test of goodness of fit developed by Professor Karl Pearson has been very roughly applied. A quantity labelled χ^2 is placed under each column of figures. If this quantity is not greater than 8, the correspondence

[1] See Chap. III, § 6 below.

between the particular distribution of expenditures and the normal curve is quite satisfactory. For values between 8 and 18, the want of agreement is not surprising. For values greater than 18, the fit is definitely bad.

These last statements are a little arbitrary, being an attempt to put a complicated, difficult and subtle measurement into non-mathematical language. The process, in more detail, has been to take ten groups of moderate size instead of the larger number given in the complete tables. These groups are, for expenditures both above and below the average,

(1) from the average to one-third of σ
(2) from one-third to two-thirds of σ
(3) from two-thirds of σ to σ
(4) from σ to five-thirds of σ
and (5) outside five-thirds of σ

where σ is the standard deviation of the expenditures in question. The value of χ^2 was then calculated from these ten groups.

Denoting by P the chance of getting, by random sampling from the normal distribution, as bad a fit as that indicated by χ^2, the values of P for ten groups are approximately:

χ^2	8	9	10	11	12	13	14	15	16	17	18	19	20
P	·53	·44	·35	·28	·24	·16	·12	·09	·07	·05	·035	·025	·018

The values of χ^2 given in the rough test of goodness of fit above are fixed by taking ·5 and ·03 as the values of P corresponding to the respective limits of expectation.

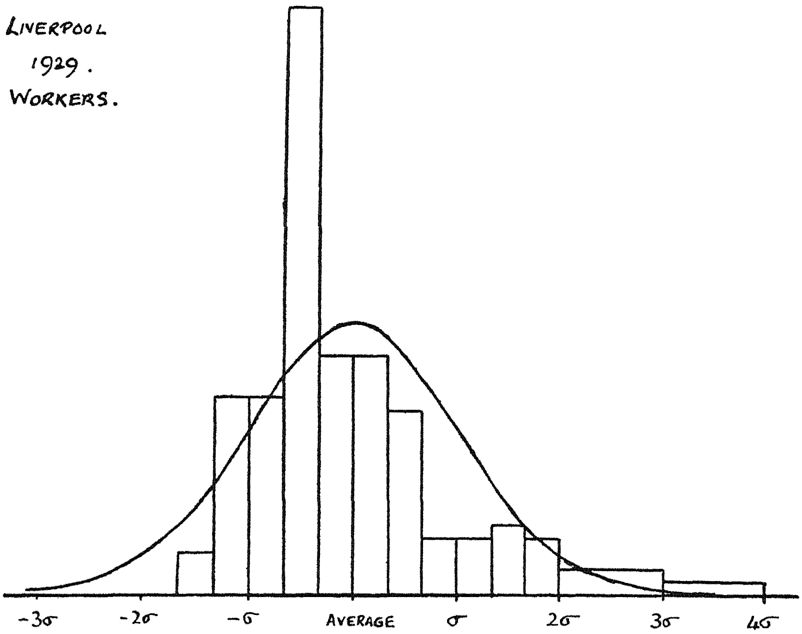
4. A Detailed Illustration of the Procedure

The method of procedure has now been adequately described and we can pass to an example set out in detail. For this purpose, we have selected the relationship between the expenditure on the vegetable food group and total expenditure in the working-class budgets collected in Liverpool in 1929.

The variation of family needs in this collection of budgets is first eliminated by calculating the number of equivalent adults in each family and then by dividing the food expenditures by this number. The distribution of

total expenditure per equivalent adult is markedly unsymmetrical. The average is 16·9 shillings per equivalent adult per week and the co-efficient of variation is 43 per cent. The distribution of expenditure per equivalent adult on the vegetable group is shown in Diagram XIV; it is also markedly unsymmetrical and a

DIAGRAM XIV.



Distribution of expenditure per equivalent adult on the vegetable food group, in comparison with the normal distribution.

large number of cases are heaped up just below the average. The average is 1·0 shillings per equivalent adult per week and the co-efficient of variation is 72 per cent.

The relation of expenditure on vegetables to total expenditure is next taken into account. The co-efficient of correlation between the two variables is found to be ·70. If y denotes the expenditure of any family on the

vegetable group and e the total expenditure, then the coefficient of regression of y on e is

$$k = r\frac{\sigma_y}{\sigma_e} = \cdot 70 \frac{72 \times 1\cdot 0}{43 \times 16\cdot 9} = \cdot 07.$$

The regression equation is thus

$$y = \cdot 07(e - 16\cdot 9) + 1\cdot 0$$

i.e. $$y = \cdot 07e + \cdot 185.$$

The original data are shown in Diagram XV, where total expenditure per equivalent adult is measured on the horizontal axis and expenditure on vegetables per equivalent adult on the vertical axis. The expenditure of each individual family is represented by a dot on this scatter diagram. The regression equation above enables us to insert the line AB to show the movement of the average expenditure on vegetables. The difference between the actual vegetable expenditure of a family and the value computed from the average equation is

$$v = y - \cdot 07e - \cdot 185$$

and is represented by a vertical distance, such as MP, which can be positive or negative.

The standard deviation of the residual expenditures v is

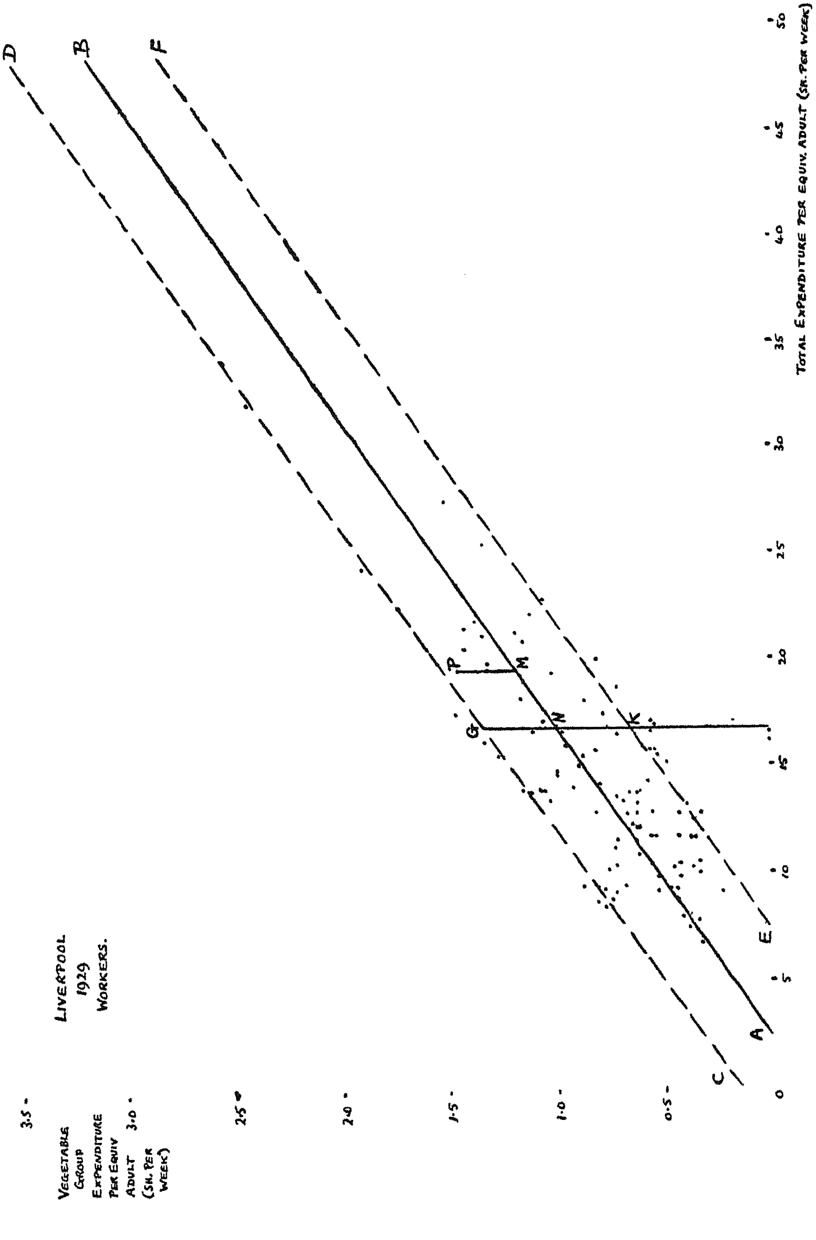
$$\sigma_v = \sigma_y\sqrt{1 - r^2} = \cdot 512$$

and the probable error of the residuals is $\cdot 674$ of $\cdot 512$, i.e. $\cdot 345$. The lines CGD and EKF, in Diagram XV, are drawn parallel to AB at vertical distances NG and KN from it, where $NG = KN = \cdot 345$. The point N marks the position of the double average.

If the distribution of the v's is normal, one quarter of the whole set of 145 observations, i.e. 36 points, would be found above the line CD, another quarter between CD and AB, between AB and EF, and below EF. Actually, the numbers are about 29, 37, 44 and 35 respectively. This variation is no more than what would be expected in a random sample of this size.

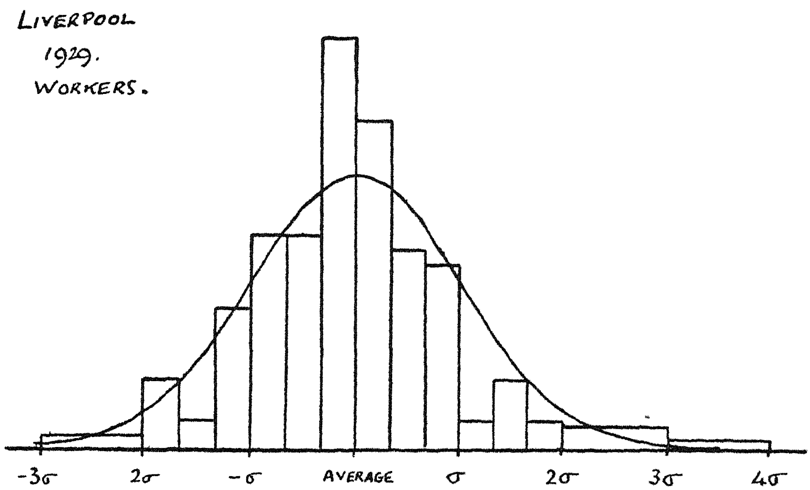
Finally, the distribution of the 145 values of v, with 79 negative values and 66 positive values, is shown in Diagram XVI. The distribution is also summarised,

DIAGRAM XV.


LIVERPOOL
1929
WORKERS.
VEGETABLE GROUP EXPENDITURE PER EQUIV ADULT (SH. PER WEEK)
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0
5
10
15
20
25
30
35
40
45
50
TOTAL EXPENDITURE PER EQUIV. ADULT (SH. PER WEEK)
D
B
F
C
A
E
P
M
G
N
K

in graded form, in the appropriate column of Table G_1 below. It is seen that the distribution is now fairly symmetrical and the correspondence with the appropriate normal curve, also shown in Diagram XVI, appears reasonably close. Using the ten grades indicated above, the value of χ^2 is computed to be roughly 14. The fit of the normal curve to the expenditure residuals, though not completely satisfactory, is by no means bad enough for the curve to be rejected as a description of the facts. Having

DIAGRAM XVI.



Distribution of residual expenditure per equivalent adult on the vegetable food group, in comparison with the normal distribution.

obtained a reasonable fit with the normal curve in this way, we have now a fairly complete account of the phenomena of the variation of tastes for the vegetable food group in this collection of budgets.

It can be noticed, from Diagram XV, that the spread of the points about the line AB does not differ greatly for successive grades of total food expenditure. One of the characteristics of the distribution of the v's that the theoretical treatment given below indicates as expected,[1] is the constancy of this spread for all ranges of total

[1] Chap. III, § 6.

expenditure. This cannot be definitely asserted here since the spread tends to increase with total expenditure. But the number of budgets is too small, and the data too imperfect,[1] for anything approaching a precise conclusion on this point. Considering this and the distributions for other expenditures and budget collections, there is some indication, at least, that this characteristic of the distribution of residual expenditures may be found approximately.

5. Examples of the Distributions of Residual Expenditures

The method now described and illustrated has been applied to the data in several budget collections. The process has usually been to compare the original distributions of expenditures on different items, or groups of items, with that given by the normal curve of error. Where the fit is bad, either because of marked want of symmetry in general or because of lack of agreement in detail, the effect of the elimination of the variation of income, or of total expenditure as a whole or on food, has been tested. Where abnormalities still remain, the elimination of the variation due to family composition has been attempted as far as possible, either by introducing the number of equivalent adults as an additional variable in the linear expenditure relation or by dividing each expenditure, family by family, by this number. Various combinations of these methods have also been used. In the case of expenditures on food items, the division by the number of equivalent adults has usually been carried out first.

In Table F, some results are shown for family expenditure on rent and on food as a whole in the cases of four budget collections. The columns headed e show the distributions of total expenditures and these are quite unsymmetrical. The columns headed y give the distributions of expenditures on rent or on food. The other columns give the distributions of residual expen-

[1] For example, the data relate to expenditure in one week only and the week selected is not the same for all families.

TABLE F. DISTRIBUTION OF FAMILY EXPENDITURE

e is total expenditure and y expenditure on a particular group. v, v^1 and v_1 are residual expenditures on a particular group: $y = ke + c + v$, $y = bn + c + v^1$ and $y = ke + bn + c + v_1$.

	Liverpool, 1929. Workers. Total*						English Towns, 1926. Clerks. Total.				L.S.E., 1932. All Classes. Total.					U.S.A., 1928–9. Farmers. Total.				
		Rent.				Normal.		Rent.		Normal.		Food.			Normal.		Food.			Normal.
	e	y	v	v^1	v_1		e	y	v		e	y	v	v_1		e	y	v	v_1	
Below -3σ	—	—	—	—	—	0·2	—	—	1	0·2	—	—	—	—	0·2	—	—	2	1	0·4
-2σ to -3σ	—	1	1	6	1	2·9	—	1	—	2·4	—	—	2	2	2·6	—	—	1	5	5·8
$-\frac{5}{3}\sigma$ to -2σ	2	—	1	4	2	3·3	1	1	2	2·8	—	2	—	—	3·1	—	1	5	2	6·7
$-\frac{4}{3}\sigma$ to $-\frac{5}{3}\sigma$	—	9	4	7	7	5·7	6	6	7	4·8	—	9	8	5	5·4	5	10	8	6	11·6
$-\sigma$ to $-\frac{4}{3}\sigma$	8	13	12	4	11	9·0	9	12	6	7·6	6	8	9	10	8·4	29	27	20	16	18·2
$-\frac{2}{3}\sigma$ to $-\sigma$	26	10	18	8	13	12·4	13	14	14	10·5	12	14	11	11	11·6	40	34	17	26	25·2
$-\frac{1}{3}\sigma$ to $-\frac{2}{3}\sigma$	18	28	21	10	20	15·6	11	16	16	13·1	37	19	22	22	14·5	44	42	47	50	31·5
Average to $-\frac{1}{3}\sigma$	25	11	22	19	21	17·4	24	12	13	14·6	29	16	11	19	16·2	35	37	52	46	35·1
Average to $\frac{1}{3}\sigma$	19	18	9	13	12	17·4	14	12	15	14·6	11	14	14	15	16·2	34	24	37	34	35·1
$\frac{1}{3}\sigma$ to $\frac{2}{3}\sigma$	13	12	14	21	15	15·6	12	14	10	13·1	11	16	14	10	14·5	26	38	23	24	31·5
$\frac{2}{3}\sigma$ to σ	7	8	7	15	6	12·4	7	5	8	10·5	5	9	13	9	11·6	18	15	15	19	25·2
σ to $\frac{4}{3}\sigma$	2	6	11	22	9	9·0	2	6	9	7·6	4	6	5	11	8·4	14	12	14	14	18·2
$\frac{4}{3}\sigma$ to $\frac{5}{3}\sigma$	5	8	4	3	8	5·7	5	6	7	4·8	1	3	8	4	5·4	9	14	10	12	11·6
$\frac{5}{3}\sigma$ to 2σ	1	4	4	1	3	3·3	3	4	1	2·8	2	4	2	2	3·1	1	7	8	2	6·7
2σ to 3σ	4	5	5	—	5	2·9	3	2	2	2·4	4	2	5	2	2·6	10	6	8	10	5·8
Above 3σ	3	—	—	—	—	0·2	2	1	1	0·2	2	2	—	2	0·2	4	2	2	2	0·4
Negative	79	72	79	75	75	$66\frac{1}{2} \pm 6$	64	62	59	56 ± 5	84	68	63	69	$62 \pm 5\frac{1}{2}$	153	151	152	152	$134\frac{1}{2} \pm 8$
Positive	54	61	54	58	58		48	50	53		40	56	61	55		116	118	117	117	
χ^2	32	25	15	22	11	—	16	11	6	—	37	9	10	10	—	33	30	30	23	—

* Total income instead of expenditure.

ditures, the residuals being obtained in different ways. The residual obtained by eliminating the effect of varying income only is denoted by v ($y = ke + c + v$); that obtained by eliminating the variation due to family composition only is denoted by v^1 ($y = bn + c + v^1$); and that obtained by eliminating both income and needs variation is denoted by v_1 ($y = ke + bn + c + v_1$). In each case, the normal distribution appropriate to the number of budgets in the collection is also given.

The distribution of total food expenditure in the London School of Economics budget collection is fairly close to normality even when no correction is applied. The symmetry is, however, improved by eliminating income and the co-efficient of variation is reduced from 34 per cent to 25 per cent. It happens, in this case, that the further correction for the number of equivalent adults does not improve the distribution; in fact, the fit to the normal distribution becomes rather worse. In contrast, the distribution of food expenditure in the budgets of farmers' families in New York State is seriously unsymmetrical whatever corrections are applied. The "lumpiness" below the average expenditure remains. It is possible that there is a good deal of variation in the consumption of food produced on the farm, and the returns would need sub-division to eliminate this influence.

The distribution of rent expenditure in the middle-class budgets collected in London and other large towns is not far removed from normality and the discrepancies are almost all removed by eliminating the income variation. The Liverpool working-class rents, on the other hand, are unsymmetrically distributed and remain so until both income and family composition variations are eliminated. Even then, there remains some irregularity, though the fit to the normal distribution is not unsatisfactory.

In the remaining Tables, the results all relate to expenditures per equivalent adult. In the course of the investigation, several plans were tried in succession. It

has seemed better to give most of the results as we obtained them, despite certain variation of method, rather than attempt to rework them on a uniform method. The computational work is, in any case, extremely heavy. Thus, for the London School of Economics budgets, we have divided the total family expenditure and its main groups by the number of equivalent adults and have then compared the grouped expenditures with the total. This method ignores the difference between expenditures on items common to the family, such as rent and fuel, and expenditures on more personal items, such as food or clothing. Again, in considering expenditures per equivalent adult on the main food groups, we have taken, for the Liverpool working-class budgets, total expenditure divided by the number of equivalent adults as the independent variable. In other collections, we have ignored non-food expenditures and compared the expenditure per equivalent adult on any food group with the total food expenditure per equivalent adult. This latter method is to be preferred.

The three sets of British budgets all display certain serious limitations and imperfections. The Liverpool working-class collection, though fairly homogeneous as regards social class, gives expenditures for one week only and the week selected varies from family to family, covering all periods of the year. A considerable fortuitous variation is to be expected, therefore, in the expenditures of this group. The middle-class budgets collected in English towns relate to employees in one occupation and can be considered as reasonably homogeneous. But the food expenditures relate to only one month of the year and some chance variation must be present. The London School of Economics budgets, on the other hand, give yearly expenditures. The budgets were, however, submitted by a self-selected group of families in response to broadcast appeals. The families are thus not homogeneous in social class and they come from various parts of the country. The collection can be taken as largely middle-class and professional but a certain num-

ber of more definitely working-class families are also included.

The German data are much more reliable. The budgets here cover working-class families in the two towns of Hamburg and Bremen and the expenditures relate to a whole year which is of practically identical extent for all families. The only imperfection likely to be serious is the fact that the families may not be typical of all working-class families in the two towns since all families selected in the first place did not return budgets. The homogeneous nature of this collection makes the smallness of the number of budgets much less of a limitation.

The results of the London School of Economics budget collection have not hitherto been published. They are brought together in Table G_3 except for those already included in Table F. The average expenditure relations obtained from this collection are to be found in the relevant tables of Chapter I.

Tables G_1, G_2 and G_3 show the distributions of expenditures on the main food groups in the three British collections. The residual expenditures, resulting from the elimination of the effect of total expenditure (either as a whole or on food only), provide a fairly good fit to the normal distribution in most cases. The main exception is the group of Miscellaneous expenditure in each case; this group, as we expect, is not amenable to treatment on the lines we have adopted.

Including the distributions of total food expenditures (taken in relation to total expenditure) with those of the food groups themselves, the three tables provide 25 distributions of residual expenditures. The value of χ^2, computed for the ten classes indicated above, is calculated in each case. The value is 8 or under in seven cases and between 9 and 18 in fourteen other cases, including three cases in which the value is as low as 9. In only five cases is the value over 18 and three of these correspond to the Miscellaneous Expenditure group. If these five cases are set aside as obeying no law, or a different law, the

TABLE G. DISTRIBUTION OF FOOD EXPENDITURE PER EQUIVALENT ADULT

e is total expenditure per equivalent adult and y expenditure per equivalent adult on a particular group. v is the residual expenditure on a particular group.

TABLE G_1. LIVERPOOL 1929 WORKERS,

The Residuals are from Regressions on Total Expenditure per Equivalent Adult

	Total. e	Total Food.		Cereals.		Meat.		Dairy Produce.		Vegetables.		Sugar, etc.		Tea, Coffee.		Miscellaneous.		Normal.
		y	v	y	v	y	v	y	v	y	v	y	v	y	v	y	v	
Below -3σ	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0·2
-2σ to -3σ	—	—	4	1	2	1	6	—	2	—	3	—	1	—	—	—	—	3·1
$-\frac{5}{3}\sigma$ to -2σ	—	2	5	—	—	—	1	—	2	—	5	—	—	—	3	—	—	3·6
$-\frac{4}{3}\sigma$ to $-\frac{5}{3}\sigma$	1	5	7	4	6	6	3	4	2	3	2	3	4	2	5	—	3	6·3
$-\sigma$ to $-\frac{4}{3}\sigma$	16	13	7	15	9	10	8	17	8	14	10	9	10	7	5	—	4	9·8
$-\frac{2}{3}\sigma$ to $-\sigma$	19	24	10	18	20	21	9	23	23	14	15	21	18	26	10	—	10	13·7
$-\frac{1}{3}\sigma$ to $-\frac{2}{3}\sigma$	26	21	17	21	22	27	20	24	21	42	15	28	19	29	35	80	40	16·9
Average to $-\frac{1}{3}\sigma$	28	23	21	21	22	17	29	15	20	17	29	16	23	21	23	25	42	18·9
Average to $\frac{1}{3}\sigma$	10	8	22	24	18	20	20	12	22	17	23	26	21	22	32	13	17	18·9
$\frac{1}{3}\sigma$ to $\frac{2}{3}\sigma$	17	13	20	10	16	8	17	14	16	13	14	18	20	14	15	4	11	16·9
$\frac{2}{3}\sigma$ to σ	7	11	13	10	12	9	12	12	8	4	13	8	15	9	3	7	7	13·7
σ to $\frac{4}{3}\sigma$	4	11	10	8	8	10	8	9	9	4	2	7	7	6	5	5	2	9·8
$\frac{4}{3}\sigma$ to $\frac{5}{3}\sigma$	6	5	3	4	2	4	4	7	6	5	5	2	2	3	2	5	2	6·3
$\frac{5}{3}\sigma$ to 2σ	1	5	—	2	1	6	3	3	1	4	2	1	—	1	3	2	2	3·6
2σ to 3σ	5	3	5	4	5	5	4	4	4	5	5	4	4	3	1	3	4	3·1
Above 3σ	5	1	1	3	2	1	1	1	1	3	2	2	1	2	3	1	1	0·2
Negative	90	88	71	80	81	82	76	83	78	90	79	77	75	85	81	105	99	$72\frac{1}{2} \pm 6$
Positive	55	57	74	65	64	63	69	62	67	55	66	68	70	60	64	40	46	
χ^2	31	21	4	15	10	25	9	22	13	58	14	27	12	35	47	Great		—

TABLE G_2. ENGLISH TOWNS, 1926. CLERKS

The Residuals are from Regressions on Total Food Expenditure per Equivalent Adult

	London and Large Towns.																Small Towns.			
	Total.	Total Food.*		Cereals.		Meat.		Dairy Produce.		Vegetables.		Sugar, etc.†	Tea, Coffee.		Miscellaneous.†	Normal.	Total.	Total Food.*		Normal.
	e.	y	v	y	v	y	v	y	v	y	v	y | v	y	v	y | v		e	y	v	
Below -3σ	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0·1	—	—	—	0·1
-2σ to -3σ	—	—	—	—	—	—	—	1	2	1	1	—	—	1	—	1·9	—	2	4	2·2
$-\frac{5}{3}\sigma$ to -2σ	—	—	1	—	1	—	1	1	1	—	3	3	—	—	—	2·3	2	2	2	2·6
$-\frac{4}{3}\sigma$ to $-\frac{5}{3}\sigma$	7	—	4	—	2	4	3	6	1	5	3	5	6	6	—	3·9	6	6	5	4·5
$-\sigma$ to $-\frac{4}{3}\sigma$	8	10	6	16	9	6	5	4	10	7	8	4	5	6	—	6·1	9	6	5	7·0
$-\frac{2}{3}\sigma$ to $-\sigma$	12	12	14	6	10	12	16	14	12	8	8	9	12	8	20	8·4	19	9	8	9·8
$-\frac{1}{3}\sigma$ to $-\frac{2}{3}\sigma$	14	14	15	12	11	15	13	13	12	19	11	14	13	17	29	10·5	19	11	14	12·2
Average to $-\frac{1}{3}\sigma$	9	18	11	16	18	12	11	13	9	16	13	15	13	11	7	11·8	8	13	14	13·6
Average to $\frac{1}{3}\sigma$	10	15	13	12	17	16	14	11	15	8	17	16	10	9	10	11·8	18	16	13	13·6
$\frac{1}{3}\sigma$ to $\frac{2}{3}\sigma$	6	4	9	10	5	11	8	9	9	7	7	8	12	11	8	10·5	14	11	8	12·2
$\frac{2}{3}\sigma$ to σ	9	6	5	7	6	2	7	4	7	5	6	4	10	11	6	8·4	5	7	17	9·8
σ to $\frac{4}{3}\sigma$	1	—	2	3	3	1	3	4	2	3	2	2	3	4	1	6·1	5	11	6	7·0
$\frac{4}{3}\sigma$ to $\frac{5}{3}\sigma$	6	1	1	3	3	3	3	2	4	4	4	3	1	1	2	3·9	4	6	3	4·5
$\frac{5}{3}\sigma$ to 2σ	6	5	4	2	2	4	4	3	1	1	1	2	—	—	—	2·3	6	3	2	2·6
2σ to 3σ	2	3	4	1	1	2	1	2	3	3	3	3	4	4	6	1·9	5	1	3	2·2
Above 3σ	—	2	1	2	2	2	1	3	2	3	3	2	1	1	1	0·1	2	—	—	0·1
Negative	50	54	51	50	51	49	49	52	47	56	47	50	49	49	56	45 ± 5	63	49	52	52 ± 5
Positive	40	36	39	40	39	41	41	38	43	34	43	40	41	41	34		59	55	52	
χ^2	17	32	20	11	15	21	14	14	7	18	8	11	12	9	Great	—	35	5	7½	—

* These residuals are from regressions on total expenditure per equivalent adult.

† The correlations are here so small that little change is found from the y to the v distributions.

TABLE G_3. L.S.E., 1932. ALL CLASSES

The Residuals are from Regressions on Total Food Expenditure per Equivalent Adult in the cases of the Food Groups, and on Total Expenditure per Equivalent Adult in Other Cases

	Total.	House-keeping.		Rent.		Cloth-ing.		Fuel.		Total Food.		Cereals.	Meat.	Dairy Pro-duce.	Vege-tables.		Sugar, etc.		Tea, Coffee.		Miscel-laneous.	Normal.
	e	y	v	y	v	y	v	y	v	y	v	y	y	y	y	v	y	v	y	v	y	
Below -3σ .	—	—	—	—	—	—	1	—	—	—	—	—	—	—	—	1	—	—	—	—	—	0·2
-2σ to -3σ	—	—	2	—	2	—	2	1	1	—	2	—	1	1	—	1	—	—	—	—	—	2·6
$-\frac{5}{3}\sigma$ to -2σ	1	4	3	—	3	1	1	3	2	3	2	3	1	3	—	1	1	4	—	2	—	3·1
$-\frac{4}{3}\sigma$ to $-\frac{5}{3}\sigma$	5	7	4	4	1	2	6	4	4	9	1	3	3	4	7	4	3	4	—	6	—	5·4
$-\sigma$ to $-\frac{4}{3}\sigma$	11	11	8	9	4	14	3	7	9	9	12	11	13	8	8	8	10	10	10	7	—	8·4
$-\frac{2}{3}\sigma$ to $-\sigma$	14	12	12	20	20	17	13	21	12	13	12	14	16	15	16	16	16	12	14	12	29	11·6
$-\frac{1}{3}\sigma$ to $-\frac{2}{3}\sigma$	17	15	18	19	24	20	15	17	15	17	23	19	14	19	21	17	21	22	29	18	23	14·5
Average to $-\frac{1}{3}\sigma$ . .	23	20	20	21	15	20	16½	12	26	19	14½	22	22	14	22	25	16	14	28	22	32	16·2
Average to $\frac{1}{3}\sigma$ . . .	17	17	15	11	15	12	27½	15	15	19	14½	16	15	18	10	14	22	22	9	18	12	16·2
$\frac{1}{3}\sigma$ to $\frac{2}{3}\sigma$ . .	9	13	13	13	13	12	15	18	14	9	14	10	12	13	16	9	7	12	14	14	9	14·5
$\frac{2}{3}\sigma$ to σ . .	11	7	12	10	8	7	11	9	9	8	11	6	7	14	9	8	10	8	5	10	6	11·6
σ to $\frac{4}{3}\sigma$ . .	3	3	8	1	6	6	5	6	3	9	10	3	8	7	3	10	8	2	5	6	5	8·4
$\frac{4}{3}\sigma$ to $\frac{5}{3}\sigma$ . .	1	6	1	8	3	3	2	2	6	2	3	10	5	3	3	2	3	4	3	3	1	5·4
$\frac{5}{3}\sigma$ to 2σ . .	5	2	2	1	3	4	3	4	4	3	2	5	2	2	4	2	1	3	2	3	—	3·1
2σ to 3σ . .	6	6	5	3	5	5	2	4	3	1	1	1	4	2	3	5	4	6	3	2	2	2·6
Above 3σ .	1	1	1	3	1	1	1	1	1	3	2	1	1	1	2	1	2	1	2	1	5	0·2
Negative . .	71	69	67	73	69	74	57½	65	69	70	66½	72	70	64	74	73	67	66	81	67	84	62±5½
Positive . .	53	55	57	50	54	50	66½	59	55	54	57½	52	54	60	50	51	57	58	43	57	40	
χ^2 . . .	24	8	4	20	24	18	15	16	11	9	10	10	9	6	23	13	16	16	40	8	Great	—

results are reasonably satisfactory as illustrating the validity of our analysis. These particular results can only establish a presumption that the analysis is applicable. More definite conclusions can only be obtained when budget collections are available with a more detailed classification of items, when more attention is paid to the variation of needs and when we have larger and more homogeneous samples of families.

A number of interesting conclusions can be tentatively drawn from the results shown in the tables. From Tables D and E above, we see that the smallest co-efficient of expenditure variation for the Liverpool working-class families occurs in the case of the Cereal group, i.e. the expenditure group of the greatest urgency. On the other hand, a very large co-efficient is found in the case of the Vegetable group and the variability is considerably reduced when the total expenditure is allowed for. Table G_1 shows that the closest fit of residual expenditures to the normal distribution is found in the case of the Meat group. Such divergences as remain are largely due to the presence of a few vegetarians in the group. The improvement in the fit in this case is very marked as we pass from the distribution of the original expenditures to that of the residual expenditures. A clear improvement is also found in the Vegetable group, a fact we have already noted and illustrated in the detailed example set out above. The other residual expenditures are only approximately normal and some skewness is to be noticed in each distribution. This might easily be due to the chance variations in weekly expenditures.

Passing to the middle-class budgets from London and large towns, rather better results are obtained. The co-efficients of expenditure variation are now smaller and the smallest of all occurs in the case of expenditure on the Dairy Produce group. Further, Table G_2 shows that the distribution of residual expenditures on Dairy Produce is a good approximation to the normal distribution and that the fit is a marked improvement on that of the distribution of original expenditures. Another case

of marked improvement is again provided by the expenditures on the Vegetable group. On the other hand, the variation of the Cereal expenditures is larger and the distribution has a definite skewness even when the effect of total food expenditure is eliminated.

In the case of the London School of Economics budget collection, which includes largely middle-class families, the co-efficient of expenditure variation is again smallest for the Dairy Produce group. From Table G_3, it is seen that all residual expenditures (except the Miscellaneous expenditures) fit the normal distribution fairly well. The expenditures on the Meat, Cereal and Dairy Produce groups are approximately normal before the effect of total food expenditure is considered. In such cases, the distributions of residual expenditures must also be normal, except in very rare circumstances, but the co-efficient of variation will be somewhat reduced. In particular, the fit of the Dairy Produce expenditure is very good. The Vegetable Expenditure group, in this as in the other collections, shows a considerably reduced co-efficient of variation, and an improved fit to the normal distribution, when the effect of total expenditure is eliminated. The fit of the distribution of expenditures on the Tea group to the normal distribution is also improved when total expenditure is allowed for.

One particular result of this comparison is almost certainly of significance. The Dairy Produce expenditures have much less variation, and fit the normal distribution more closely, in the cases of middle-class families than in the working-class case. This fact supports the conclusion already established that Dairy Produce tends to be an expenditure of greater urgency for middle-class families than for working-class families.

The validity of our analysis can be subjected to a rather more searching test by considering the distributions of expenditures on particular food items rather than on broad groups of items. Table H shows expenditure distributions for three selected items in the case of the Liverpool workers and for ten selected items in the case

TABLE H. DISTRIBUTION OF FOOD EXPENDITURE PER EQUIVALENT ADULT

The Residuals v are from Regressions on Total Food Expenditure per Equivalent Adult

	Liverpool, 1929. Workers.								Hamburg and Bremen, 1927–8. Workers.																				
	Total Food.	Bacon.		Bread.		Tea.		Normal.	Total Food.	Rye Bread.		Wheat Bread.		Milk.		Butter.		Mar-garine.		Po-tatoes.		Vege-tables.		Fresh Fruit.		Sug-ar.*	Coffee.		Normal.
		y	v	y	v	y	v			y	v	y	v	y	v	y	v	y	v	y	v	y	v	y	v	y	y	v	
Below -3σ	—	—	—	—	—	—	—	0·2	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	0·1
-2σ to -3σ	—	—	4	—	—	3	4	3·1	1	1	1	—	—	1	3	—	1	—	—	2	2	—	—	—	—	—	—	—	2·2
$-\frac{5}{3}\sigma$ to -2σ	2	—	5	2	3	—	2	3·6	3	7	2	—	2	3	4	—	3	5	3	1	2	—	1	—	1	1	—	2	2·6
$-\frac{4}{3}\sigma$ to $-\frac{5}{3}\sigma$	5	—	2	8	7	3	3	6·3	6	3	3	5	4	6	2	—	2	6	7	5	5	6	4	3	4	6	2	3	4·5
$-\sigma$ to $-\frac{4}{3}\sigma$	13	—	5	14	12	10	11	9·8	5	7	13	15	9	8	7	22	5	5	8	5	8	7	8	10	8	12	8	6	7·0
$-\frac{2}{3}\sigma$ to $-\sigma$	24	43	13	6	9	24	13	13·7	8	10	14½	14	19	12	12	12	13	8	8	19	12	17	9	16	12	8	17	13	9·8
$-\frac{1}{3}\sigma$ to $-\frac{2}{3}\sigma$	21	13	18	22	25	21	24	16·9	17	14	9½	10	10	11	10	20	15	16	16	14	10	13	14	14	17	15	18	18	12·2
Average to $-\frac{1}{3}\sigma$	23	23	31	27	23	21	21	18·9	17	6	11	14	10	12	14	16	21	17	13	13	16	16	22	16	19	15	17	19	13·6
Average to $\frac{1}{3}\sigma$	8	30	22	25	25	12	21	18·9	10	14	9	14	13	12	14	5	16	11	12	7	13	11	18	13	11	15	17	12	13·6
$\frac{1}{3}\sigma$ to $\frac{2}{3}\sigma$	13	7	16	10	13	16	15	16·9	15	11	12½	6	11	11	9	4	7	6	9	9	10	10	12	12	13	5	8	10	12·2
$\frac{2}{3}\sigma$ to σ	11	12	10	10	9	15	13	13·7	7	13	12	7	8	9	13	10	6	15	11	15	10	8	3	9	8	10	3	9	9·8
σ to $\frac{4}{3}\sigma$	11	2	8	8	6	5	3	9·8	3	11	4½	6	4	10	7	3	2	6	6½	2	4	5	4	3	4	6	4	6	7·0
$\frac{4}{3}\sigma$ to $\frac{5}{3}\sigma$	5	3	4	5	4	5	6	6·3	7	3	8	5	6	2	5	2	5	1	3½	4	5	5	5	2	2	5	2	—	4·5
$\frac{5}{3}\sigma$ to 2σ	5	6	3	8	4	4	5	3·6	2	3	1	2	5	6	1	2	2	4	2	4	4	2	—	2	1	2	4	2	2·6
2σ to 3σ	3	5	3	—	2	4	3	3·1	2	1	3	6	3	1	3	5	5	4	5	4	3	2	1	2	1	4	2	2	2·2
Above 3σ	1	1	1	—	3	2	1	0·2	1	—	—	—	—	—	—	3	1	—	—	—	—	2	3	2	3	—	2	2	0·1
Negative	88	79	78	79	79	82	78	72½ ± 6	57	56	54	58	54	53	52	70	60	57	55	59	55	59	58	59	61	57	62	61	52 ± 5
Positive	57	66	67	66	66	63	67		47	48	50	46	50	51	52	34	44	47	49	45	49	45	46	45	43	47	42	43	
χ^2	21	110	16	21	18	17	7	—	7	8½	8	19	14	3	4	40	14	12	6	21	3½	13	16	14	11	12½	26	12	—

* The correlation here is so small that little change is found from the y to the v distributions.

of the workers in Hamburg and Bremen. Considering the small number of families included in each case, the fit of the residual expenditures, after the elimination of the effect of total food expenditure, to the normal distribution is surprisingly good in almost every instance. The homogeneous German data provide particularly satisfactory results. The value of χ^2, computed for ten classes as before, is 8 or under in five cases and between 9 and 18 in the remaining eight cases. There is, therefore, no instance of a definitely bad fit.

The co-efficient of expenditure variation is small for Rye Bread and Potatoes, and very little larger for Sugar and Milk, in the German data. On the other hand, the variability of expenditures on Butter (in the German data) and on Bacon (in the Liverpool data) is very high.

Considering the actual distributions shown in Table H, the cases of Milk and Potatoes, for which there are no clear "substitutes," stand out unmistakably. In both cases, the distribution of residual expenditures is symmetrical and very close to normality. The original expenditures on Milk also fit the normal distribution very closely but the original expenditures on Potatoes form a much less even distribution.

The expenditures on Butter, before attention is paid to total food expenditure, are distributed in a highly unsymmetrical way. The symmetry and the fit to the normal distribution are considerably improved, however, when we pass to the residual expenditures. The latter still display some skewness, with a "lumpiness" just below the average and a considerable "tail" of high expenditures. It might be expected that Margarine expenditures would show a skewness in the opposite direction in compensation. This is not the case. The distribution of residual expenditures on Margarine is quite close to normality and little skewness is evident. There is possibly less connection between Butter and Margarine expenditures than might at first be thought likely. Rather similar remarks apply to expenditures on Wheat and Rye Bread. Residual expenditures on Wheat Bread display a certain

skewness whereas the corresponding expenditures on Rye Bread fit the normal curve quite well. It appears that the hypothesis of normal distribution of tastes is applicable to the "inferior" of these pairs of items but is more doubtful in its application to the "superior" of the items.

The distributions of residual expenditures on Vegetables (other than potatoes) and on Fresh Fruit also display certain skewness in the German data. This is, however, largely due to the presence in the group of families of three families of vegetarian inclination. The variability of these two expenditures is more definitely reduced than in the other cases when we pass from the original expenditures to the expenditures from which the effect of total food expenditure has been eliminated.

One further point has been examined in connection with expenditures on such "substitute" items as Wheat and Rye Bread or Butter and Margarine. It might be expected that an expenditure of this kind would not conform to the normal distribution but would show relatively fewer cases near the average and relatively more cases away from the average. This would indicate "swings" of tastes from family to family, some families definitely preferring (say) Wheat to Rye Bread and others definitely preferring Rye to Wheat Bread, after due allowance has been made for variation of total expenditure. The distribution of expenditure on Rye Bread shows some slight evidence of this characteristic. But, on the whole, it seems more justifiable to take the hypothesis of normal distribution of tastes, with a concentration of expenditures about the average. Average tastes and average expenditures do not appear to be fictitious concepts even in cases of "substitute" items of expenditure.

6. Correlation between Expenditures on Different Foods

The points just considered introduce the interesting questions of whether two items of expenditure are sub-

stitutes or complements, and of the measure of how far the items are substitutable one for the other or how far they are complementary. It would be tempting to explore this question with the aid of the budget data we have available but, in fact, it is easily seen that we can only go so far as to say that variation of preferences for two items are correlated or not over the group of families considered. The data do not allow us to distinguish the case where a family, on account of its own special tastes, takes more than the average of butter and less than the average of margarine from the special relation that these are complete substitutes one for the other. Nor do the data allow us to distinguish the case where a family has more than the average inclination for both sugar and tea from the special relation that 1 lb. of sugar is needed to sweeten $\frac{1}{4}$ lb. of tea. Definite results can only be obtained in the extreme case where it happens that there are no uses for the items other than as complements to one another in fixed proportions or as perfect substitutes.

The simple correlation co-efficient between expenditures on two items does not, by itself, give any information on such questions. In nearly all cases, the higher the income the greater is the consumption of each item and the correlation co-efficients have positive values simply for this reason. In the rare cases where the value of the constant k is negative for one item of the pair, the correlation co-efficient is, of course, negative.

The only method of eliminating the influence of income in this respect is that of partial regression. The examples we give do not relate to total income or expenditure, but to total expenditure on food per equivalent adult and to the connection between this and expenditures per equivalent adult on particular food items. The partial regression equation is, therefore, written in the following way.

Write f for the total food expenditure per equivalent adult and m and d for expenditures per equivalent adult on (say) the Meat group and on the Dairy Produce group.

In each case, the expenditure is measured from its average over the whole group of families. Then

$$m = R_{md}\sigma_1 d + R_{mf}\sigma_2 f$$

where

$$R_{md} = \frac{r_{md} - r_{mf}r_{df}}{\sqrt{(1 - r^2_{mf})(1 - r^2_{df})}}$$

$$R_{mf} = \frac{r_{mf} - r_{md}r_{df}}{\sqrt{(1 - r^2_{md})(1 - r^2_{df})}}$$

$$\sigma_1 = \frac{\sigma_m}{\sigma_d}\sqrt{\frac{1 - r^2_{mf}}{1 - r^2_{df}}} \quad \text{and} \quad \sigma_2 = \frac{\sigma_m}{\sigma_f}\sqrt{\frac{1 - r^2_{md}}{1 - r^2_{df}}}.$$

In this relation, σ_f, σ_m and σ_d denote the standard deviations of the expenditures f, m and d, and the r's denote the simple correlation co-efficients between the two expenditures indicated by the suffixes.

R_{md} is the correlation co-efficient between Meat and Dairy Produce expenditures when the influence of total food expenditure is eliminated. Its value cannot exceed the limits $+1$ and -1. For any given value of f the *average* increase of m is the average increase of d multiplied by the factor $R_{md}\sigma_1$. The actual increases of m vary, of course, from family to family in the group.

If Meat and Dairy Produce are perfectly complementary and have no other uses, then $r_{md} = +1$ and therefore $r_{mf} = r_{df}$. Hence $R_{md} = +1$ and the regression equation reduces to

$$m = \frac{\sigma_m}{\sigma_d} d$$

i.e. m varies exactly as d. If the items are perfect substitutes and have no other uses, then $r_{md} = -1$ and $r_{mf} = -r_{df}$. Hence $R_{md} = -1$ and the equation becomes

$$m = -\frac{\sigma_m}{\sigma_d} d$$

i.e. m varies exactly as d but in the opposite sense. These are the two extreme cases of which, of course, we have found no examples in the data examined.

If $r_{md} = r_{mf}r_{df}$, so that $R_{md} = 0$, then the correlation between the two expenditures is due solely to their separate relationships to the total food expenditure.[1]

If $r_{md} > r_{mf}r_{df}$, so that $R_{md} > 0$, then there is a positive correlation between preferences for *m* and *d* over the families of the group. This may be due *either* because (for example) the same family prefers to spend more than the average on both beef and cheese, *or* because the persons like to eat cheese after beef.

If $r_{md} < r_{mf}r_{df}$, so that $R_{md} < 0$, there is a negative correlation between preferences for *m* and *d* over the families, whether because some families tend to be vegetarian and others carnivorous or because beef and cheese are alternative forms of nourishment.

In spite of this ambiguity of the causes of the values of R_{md}, it is useful to examine some actual cases. The correlations between the expenditures on the main food groups have been worked out for two sets of budgets and the results are exhibited in Table I. The first co-efficients (r_{xy}) represent the crude correlations between the expenditure groups and total food expenditure, and between the various groups. The second co-efficients (R_{xy}) represent the partial, or corrected, correlations between the various food groups.

Thus, if we take Bread and the Meat group in the London School of Economics budget collection, we have the crude correlations

$$r_{mf} = \cdot 753\ ;\ r_{bf} = \cdot 442 \quad \text{and} \quad r_{mb} = \cdot 229.$$

Hence, $\quad r_{mb} - r_{mf}r_{bf} = \cdot 229 - \cdot 333 = - \cdot 104$

$$\sqrt{1 - r^2_{mf}} = \cdot 658 \quad \text{and} \quad \sqrt{1 - r^2_{bf}} = \cdot 897.$$

So $\quad R_{mb} = - \cdot 104 / \cdot 658 \times \cdot 897 = - \cdot 18.$

The correlation between bread and meat (r_{mb}) is only positive because both expenditures increase with food expenditure. For constant food expenditure, the correlation (R_{mb}) is negative.

[1] It may be noticed that the case where $r_{md} = 0$ does not lead to any important condition.

TABLE I. CORRELATIONS BETWEEN FOOD EXPENDITURES PER EQUIVALENT ADULT

Food Groups.	Liverpool, 1929. Workers.						L.S.E., 1932. All Classes.					
	Total Food.	Bread.	Meat.	Dairy Produce.	Vegetables.	Sugar, etc.	Total Food.	Bread.	Meat.	Dairy Produce.	Vegetables.	Sugar, etc.
r_{xy}.												
Bread	—·06						·44					
Meat	·81	—·17					·75	·23				
Dairy Produce	·80	—·02	·48				·65	·08	·25			
Vegetables	·84	·14	·56	·57			·74	·31	·30	·46		
Sugar, etc.	·44	·24	·23	·29	·38		·44	·12	·31	·15	·28	
Tea, Coffee	·59	·16	·42	·44	·42	·25	·67	·21	·46	·26	·40	·41
R_{xy}.												
Meat		—·20						—·18				
Dairy Produce		·05	—·46					—·30	—·48			
Vegetables		·34	—·26	—·25				—·03	—·57	—·03		
Sugar, etc.		·30	—·23	—·12	·02			—·09	—·04	—·20	—·07	
Tea, Coffee		·20	—·06	—·04	—·05	—·01		—·13	—·09	—·31	—·18	·17

Hamburg and Bremen, 1927–8. Workers

r_{xy}.	Total Food.	Milk.	Butter.
Milk	·49		
Butter	·68	·48	
Margarine	— ·17	— ·33	— ·80

r_{xy}.	Total Food.	Potatoes.	Vegetables.
Potatoes	·21		
Vegetables	·77	·12	
Fresh Fruit	·42	— ·17	·74

r_{xy}.	Total Food.	Sugar.
Sugar	·11	
Coffee	·41	— ·13

R_{xy}.	Milk.	Butter.
Butter	·22	
Margarine	— ·29	— ·94

R_{xy}.	Potatoes.	Vegetables.
Vegetables	— ·06	
Fresh Fruit	— ·29	·73

R_{xy}.
Sugar and Coffee — ·19

Amsterdam, 1923–4. All Classes

r_{xy}.	Total Food.	Sugar.
Sugar	·60	
Coffee	·32	·285

R_{xy}.
Sugar and Coffee ·14

All the crude correlations listed are, in fact, positive with the exception of three correlations involving bread expenditure in the Liverpool budgets. The linear relation between bread and total food expenditure is found to have a negative value of k, i.e. bread expenditure decreases with increasing total expenditure, in the group of Liverpool families. The correlation between bread and total food expenditure is necessarily negative and this gives rise also to the negative correlations between bread expenditure and two of the other expenditure groups.

The corrected correlations, on the other hand, are usually negative. This is what we expect from the interpretation given above. The only significant exceptions occur in the relations between bread expenditure and other groups of expenditures in the Liverpool budgets. It is here fairly safe to assume that, at the income levels of the workers in Liverpool, bread is complementary in its relations to most other items of expenditure such as vegetables, sugar and tea. Since bread is an item of great urgency in this group of families, this conclusion is in line with the theoretical deduction that necessary items of expenditure must be in considerable complementary relationship with other items.[1] It must be stressed, however, that positive values of the corrected correlations do not indicate, at all necessarily, that the items are complementary. This has been clearly pointed out above.

Some of the corrected correlations have values which are not significant. It is clear, for example, that there is little correlation between expenditures on either sugar or tea and other groups of expenditure (when the effect of total expenditure is allowed for) in the London School of Economics budgets. There is a possibly significant *positive* corrected correlation between sugar and tea expenditures. Again, in the Liverpool working-class budgets, there is practically no correlation between tea expenditure and other expenditure groups.

[1] See Chap. III, § 8, below.

Table I also gives a few correlations that have been obtained for selected food items in the working-class budgets collected in Hamburg and Bremen (1927–8). These correlations are almost all significant. There is a positive corrected correlation between milk and butter and between vegetables and fresh fruit. On the other hand, the corrected correlation co-efficient between butter and margarine is negative and numerically nearly equal to unity. The inverse relationship between expenditures on butter and margarine, amongst families at a given level of total food expenditure, is almost a perfect one. The correlations between potatoes and vegetables and between potatoes and fresh fruit are also negative.

In the Amsterdam budget collection (1923–4), the corrected correlation between sugar and coffee is positive and probably significant. This is in agreement with the corresponding corrected correlation in the London School of Economics budgets. Both these collections include a considerable proportion of middle-class families. But, in the case of the purely working-class budgets of Hamburg and Bremen, there is a negative corrected correlation between sugar and coffee. The corresponding correlation for the Liverpool workers, though negative, is too small to be significant. It does appear, however, that the relation between expenditures on sugar and on tea or coffee may be different for workers than for middle-class families.

The nature of the corrected correlation co-efficient R_{md} may be further illustrated by developing an alternative expression of the relation between the two expenditures m and d. Write the average linear relations between each of the expenditures m and d and the total food expenditure f, all measured from their average values, in the form

$$m = r_{mf}\frac{\sigma_m}{\sigma_f}f \qquad \text{and} \qquad d = r_{df}\frac{\sigma_d}{\sigma_f}f.$$

Let v_m and v_d denote the residual expenditures for any family, i.e. the differences between the actual m and d for the family

and the average values from the relations above. Then, for any family,

$$m = r_{mf}\frac{\sigma_m}{\sigma_f}f + v_m \qquad \text{and} \qquad d = r_{df}\frac{\sigma_d}{\sigma_f}f + v_d.$$

It follows quite easily that

$$\begin{aligned}\text{Mean }(v_m) = \text{Mean }(v_d) &= \text{Mean }(v_m v_d) = \text{Mean }(v_m f)\\ &= \text{Mean }(v_d f) = 0\end{aligned}$$

$$\text{Mean }(mv_m) = \text{Mean }(v_m^2)$$
$$\text{Mean }(dv_d) = \text{Mean }(v_d^2)$$

$$\text{Mean }(m^2) = r_{mf}\frac{\sigma_m}{\sigma_f}\text{ Mean }(mf) + \text{Mean }(mv_m)$$

and $$\text{Mean }(d^2) = r_{df}\frac{\sigma_d}{\sigma_f}\text{ Mean }(df) + \text{Mean }(dv_d).$$

Hence,

$$\sigma_{v_m} = \sigma_m\sqrt{1 - r_{mf}^2} \qquad \text{and} \qquad \sigma_{v_d} = \sigma_d\sqrt{1 - r_{df}^2}$$

as is otherwise known.

Further,

$$\begin{aligned}\text{Mean }(v_m v_d) &= \text{Mean}\{(r_{mf}\frac{\sigma_m}{\sigma_f}f - m)(r_{df}\frac{\sigma_d}{\sigma_f}f - d)\}\\ &= \sigma_m\sigma_d(r_{mf}r_{df} - r_{mf}r_{df} - r_{mf}r_{df} + r_{md})\\ &= \sigma_m\sigma_d(r_{md} - r_{mf}r_{df})\end{aligned}$$

So

$$r_{v_m v_d} = \frac{\text{Mean}(v_m v_d)}{\sigma_{v_m}\sigma_{v_d}} = \frac{r_{md} - r_{mf}r_{df}}{\sqrt{\{(1 - r^2_{mf})(1 - r^2_{df})\}}} = R_{md}$$

R_{md} is, therefore, the co-efficient of correlation between the residual expenditures on the two groups after allowance has been made for the effect of total food expenditure. In other words, when the data are classified in groups of equal expenditure, R_{md} is the correlation co-efficient between the variations of the m and d expenditures from the average values.

CHAPTER III

THEORETICAL ANALYSIS OF THE DISTRIBUTION OF EXPENDITURE

1. Preference Scales and Indifference Curves

The problem to be analysed can be first expressed in quite general terms. An individual consumer has, in a definite period of time, a given sum of money to spend on various consumers' goods and services obtainable at given market prices. In what way will the individual distribute his total expenditure? Further, how is the distribution of expenditures on different items modified when the total sum to be spent changes, or when the relative market prices change? We have here a problem of general equilibrium. The whole complex of consumers' goods and services, their market prices and the expenditures devoted to them, is considered and the effect of any change in the price and income structure is to be traced. A solution of this general problem is essential; without it we cannot proceed to more particular studies in which certain aspects of the distribution of expenditure are isolated for detailed examination.

The analysis of our problem is limited, for the moment, to the simplest case where only two goods enter into the individual's calculations. We take as our starting point the assumption that the distribution of the individual's expenditure is based on a definite scale of preferences. It is assumed, in fact, that the whole set of possible combinations of purchases of the two goods can be arranged by the individual in order of ascending preference. Faced with two alternative combinations of purchases, the individual can either distinguish the preferable combina-

tion or classify them as indifferent, at the same level of preference. Notice that the individual's preferences are only ordered according to the qualitative criterion of greater or less; there is no question of assuming a quantitative measure of preference, and statements such as that one preference is (say) twice or three times another are not needed. Further, no particular psychological implication is involved in the assumption of a scale of preferences. It is not assumed that the individual acts consciously or "rationally," basing his choices on any "standard of value." All that we want is the knowledge that the distribution of the individuals' expenditure can, but not must, be accounted for by means of a definite preference scale of some form.[1]

The preference scale of the individual can be represented most easily in diagrammatic terms. Selecting two perpendicular axes in a plane, we measure the purchases of one good X_1 on the horizontal axis Ox_1 and the purchases of the other good X_2 on the vertical axis Ox_2. Any point P in the plane (above Ox_1 and to the right of Ox_2) then represents a possible combination of purchases, the purchase of X_1 being read off as OM along the horizontal axis and the purchase of X_2 as ON along the vertical axis (Fig. 1). As P moves in the plane, the purchases vary in a way that is clearly seen from the diagram.

Starting from any combination of purchases represented by a point P, the individual is able to select, on our assumption, all those other combinations of purchases which are indifferent in comparison with the given combination. These combinations are represented by points lying on some curve I_1 passing through P, i.e. on an *indifference curve*. The purchases represented by any point on I_1 are not preferred to those represented by any other point on the curve. Next, taking a combination of purchases represented by a point Q not on I_1, a second indifference curve I_2 can be drawn to pass

[1] For an account of the relationship between economics and psychology, see Robbins, *The Nature and Significance of Economic Science*, pp. 83–6.

through Q and to include all points representing purchases indifferent in comparison with those of Q. This process can be repeated, and the set of indifference curves extended, until all possible combinations of purchases have been accounted for. Finally, the indifference curves so constructed can be arranged, according to the scale of preferences, in ascending order of preference as $I_1, I_2, I_3, \ldots$

The complete set of indifference curves suffices to represent the individual's preference scale. If one com-

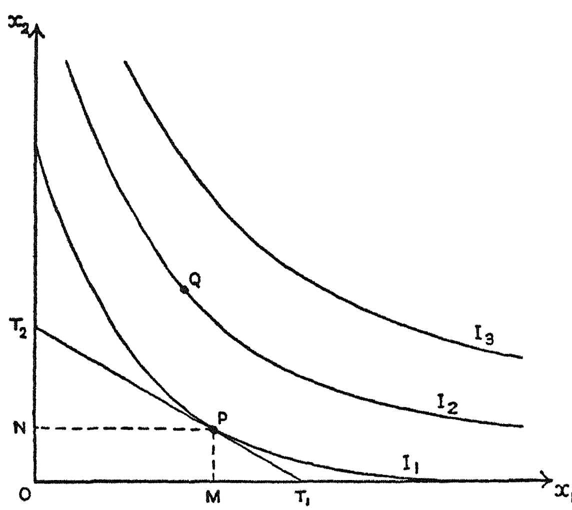


FIG. 1.

bination of purchases is represented by a point on a higher indifference curve than a second combination, then the first purchases are preferred to the second. If the points are on the same indifference curve, then the individual can express no preference for the one combination of purchases over the other; the purchases are at the same level of preference.

For simplicity of treatment, it is taken that the preference scale and the indifference curves of the individual are continuous. This implies, firstly, that the purchases of

the individual can vary by indefinitely small amounts and, secondly, that there is a continuous variation from one indifferent combination of purchases to others. There is now an indefinitely large number of indifference curves covering the part of the plane marked off by the axes and each curve is smooth and continuous. The curves are still arranged in ascending order of preference; the order is continuous and includes an indefinitely large number of curves. It must be remembered, therefore, that the following development makes no allowance for possible discontinuities in the preference scale.

Next, it is taken that no two indifference curves intersect and that the level of preference increases as we move in the plane upwards and to the right away from *O*. This implies that we are limiting our considerations to an individual under conditions where an increase in the purchase of one good, unaccompanied by a decrease in the other, increases the level of satisfaction. It follows that the level of preference can only remain unaltered if an increase in the purchase of one good is offset by some decrease in the purchase of the other. In other words, the indifference curves must slope downwards to the right at all points. Finally, it is taken that each indifference curve is convex to the origin *O* at every point. The significance of this last condition is described below.

A set of indifference curves possessing all these properties is said to be of "normal" form and the curves appear as shown in Fig. 1. It should be noticed that there is no element of necessity about the normal form of the indifference curves. The analysis can proceed if the curves assume more complicated forms; the results are then rather more involved and are subject to certain inconvenient limitations. It is found, however, that all observable market phenomena are adequately accounted for by the normal indifference curves and this is the only reason for limiting our analysis to such curves. The normal form of the curves certainly ceases to be appropriate when the individual purchases the goods in such large quantities that they become burdensome or when

his purchases leave him below the "subsistence" level. But such situations are outside any theory designed to apply to market conditions.

Consider any combination of purchases represented by a point P in the plane Ox_1x_2. A tangent line can be drawn to the indifference curve at P intersecting the axes at T_1 and T_2 (Fig. 1). The indifference curves being of normal form, this line must slope downwards to the right and its gradient is negative. The numerical value of the gradient (referred to the horizontal axis) is measured by the ratio of OT_2 to OT_1.

In order to preserve the level of preference represented by the purchases of P, a small decrease in the purchase of X_1 must be compensated by a corresponding increase in the purchase of X_2. This change is represented by a move along the indifference curve to the left from P and, in the limit for very small changes, by a move along the tangent line to the left from P. The numerical gradient of the tangent line is thus the ratio of the small increase in X_2 to the small decrease in X_1 which compensate each other and leave the level of preference unaltered. This ratio may be termed the *marginal rate of substitution* of the good X_2 for the good X_1.[1]

The normal form of the indifference curves requires that the curve at P should be convex to the origin. As we move along the indifference curve to the left from P, the numerical gradient of the tangent line to the curve increases since OT_2 increases and OT_1 decreases. Hence, as the individual continues to substitute the good X_2 for the good X_1 (at the same level of preference), the marginal rate of substitution of X_2 for X_1 increases. It requires continuously more of X_2 to compensate for a given small loss of X_1 as the substitution goes on. This is what can be called the *principle of increasing marginal rate of substitu-*

[1] This term was introduced by Hicks and Allen, "A Reconsideration of the Theory of Value," *Economica*, 1934, pp. 55 *et seq.* The marginal rate of substitution is equivalent, in the terminology of Pareto, to the ratio of the marginal utility of the good X_1 to that of the good X_2. It is found preferable, for reasons stated in the article quoted, to avoid the use of the marginal utility concept here.

tion and it expresses, in a convenient way, the condition that the indifference curves are convex to the origin.

2. The Equilibrium Distribution of Expenditure between Two Goods

To solve our general problem, we now suppose that the individual has a fixed sum of money e to spend on the two goods X_1 and X_2 which he can obtain at fixed market prices p_1 and p_2. It is required to find that combination of purchases which will take him as far up his scale of preferences as is consistent with the assumed conditions of fixed expenditure and prices. The solution of this problem can be presented in diagrammatic terms in the following way.

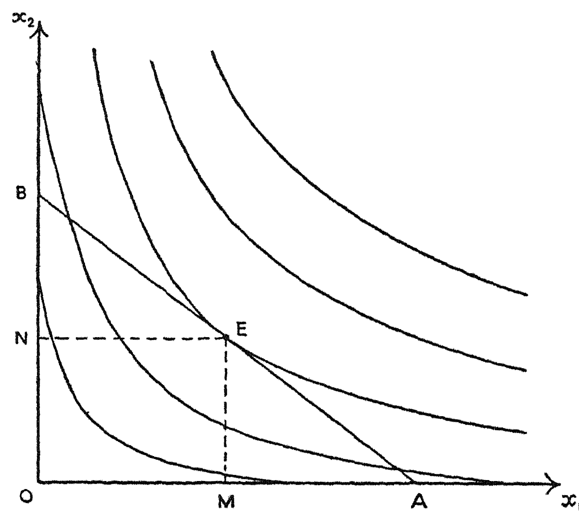


Fig. 2.

Whatever purchases the individual makes, the sum of the expenditures on X_1 and X_2 must equal the given total expenditure e. If the whole expenditure is devoted to X_1, an amount e/p_1 can be purchased at the price p_1. Represent this by the length OA on the horizontal axis Ox_1 (Fig. 2). If the whole expenditure is devoted to X_2,

an amount e/p_2 represented by OB on the axis Ox_2 can be purchased. If the whole expenditure is divided between the two goods, then the purchases x_1 and x_2 must be represented by some point between A and B on the line AB, since the equation of the line is

$$p_1x_1 + p_2x_2 = e.$$

For example, if a third of the expenditure is devoted to X_1 and two-thirds to X_2, then the purchases are $\frac{1}{3}OA$ and $\frac{2}{3}OB$ and these are clearly represented by a point on the line AB. Hence, the possible purchases of the individual, consistent with the given total expenditure and prices, all lie on the line AB. The gradient of AB is negative and its numerical value is the ratio of OB to OA, or of e/p_2 to e/p_1, i.e. of p_1 to p_2. The given price ratio thus measures the numerical gradient of AB.

It is evident that the line AB cuts many of the indifference curves and touches one of them at the point E. The indifference curve touched by AB is at a higher level of preference than any of those cut by the line. This is as far as the individual can go up his scale of preferences with the given total expenditure and the given prices. Hence, the individual cannot do better than make the purchases represented by the point E, buying an amount OM of X_1 and an amount ON of X_2.

The implications of the normal form of the indifference curves are now clear. Since the indifference curves slope downwards to the right, it is *possible* for the line AB to touch one of them. Since the curves are convex and non-intersecting, the line AB can touch *only one* curve. Finally, since the level of preference rises as we move upwards and to the right, the level of preference at E is higher than on any indifference curve cut by AB. The point E is thus uniquely determined with a level of preference as great as possible under the conditions. The normal form of the indifference curves makes for uniqueness and for stability in the distribution of a given expenditure at given prices. If the indifference curves assume other forms, there may be two or more points

such as E and some of them would not give stable distributions of expenditure.

The given line AB is a tangent to the indifference curve at the point E. Hence, the ratio of p_1 to p_2 (the gradient of AB) is equal to the marginal rate of substitution of X_2 for X_1 (the gradient of the indifference curve) at the point E. This relation is fundamental and enables us to solve the problem of the distribution of expenditure on two goods.

The purchases represented by E can be called the *equilibrium purchases* of the individual. These purchases are determined by the two conditions:

(1) The marginal rate of substitution of X_2 for X_1 is equal to the given ratio of the prices of X_1 and X_2.

(2) The separate expenditures on X_1 and X_2 add to the given total expenditure.

Further, for indifference curves of normal form, the determination of the equilibrium purchases is unique and the distribution of expenditure is a stable one.[1]

The position of the line AB, and hence of the point E, changes when the individual has a different total sum to spend or when the market prices are different. The equilibrium purchases, or the demands of the individual for the two goods, are uniquely determined for each total expenditure and each pair of prices but change when the expenditure or prices change. Our next problem is to trace the variation of the equilibrium purchases consequent upon definite changes in total expenditure or in market prices. We can confine our attention here to the first kind of change, i.e. a change of total expenditure prices remaining fixed.

Since the gradient of the line AB is numerically equal to the price ratio, the line moves parallel to itself and

[1] The use of the term "equilibrium" must not be taken as implying that there is a tendency on the market towards the realisation of the equilibrium purchases. All we have shown is that the equilibrium purchases are the only ones consistent with a maximum level of preference for the given expenditure and prices.

away from O as total expenditure increases. Fig. 3 shows a series A_1B_1, A_2B_2, A_3B_3, . . . of positions of the line AB for increasing total expenditure. We have then a series of points E_1, E_2, E_3, . . . where indifference curves are touched. If total expenditure is allowed to increase gradually and continuously, the points will lie on a continuous curve across the indifference curves. The varying co-ordinates of a point on this curve give

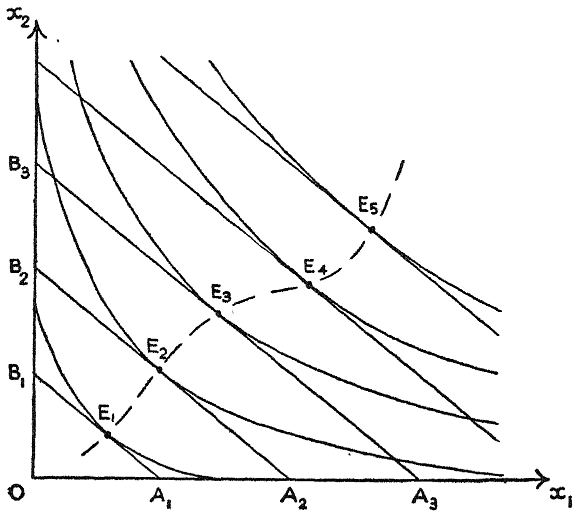


FIG. 3.

the variation of the equilibrium purchases as total expenditure increases.

The shape of the curve cannot be determined unless we know the form of the indifference curves.[1] We may expect that the curve will slope upwards to the right as shown in Fig. 3, in which case the purchases and expenditures on both goods increase with total expenditure. Even when the indifference curves are of normal form, however, it is quite possible that the curve will slope upwards to the left or downwards to the right.

[1] One case is instanced in § 4 below, where the curve reduces to a straight line.

H

Fig. 4 illustrates this possibility. In this case, which may be taken as exceptional, the purchase of, and expenditure on, one of the goods decreases as the total expenditure increases.[1]

3. The General Case of Three or More Goods

The fundamental assumption of the existence of a preference scale applies no matter how many goods are

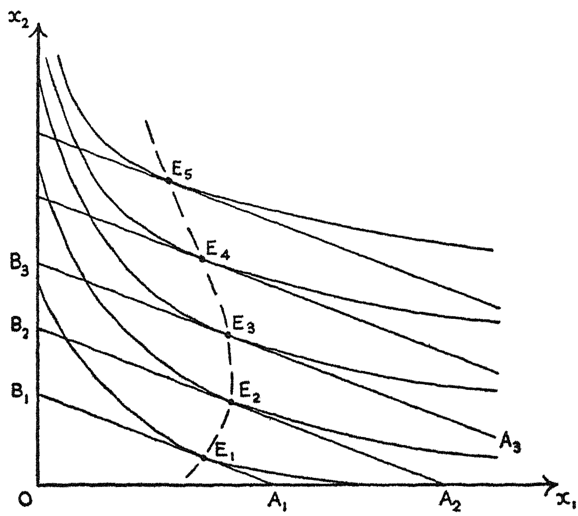


Fig. 4.

considered by the individual. The various possible combinations of purchases of the goods can be ordered according to increasing preference exactly as before. The diagrammatic representation, however, becomes more complicated when there are three goods and finally impossible, at least in terms of physical models, when there are more than three goods.

The representation, in the case of three goods X_1, X_2

[1] The path of the point E and the variation of demand as total expenditure increases are more fully described in Hicks and Allen, *op. cit.*, pp. 61 *et seq.*

and X_3, needs a three-dimensional construction. Taking three mutually perpendicular axes, Ox_1, Ox_2 and Ox_3, for the measurement of the purchases of the three goods, any combination of purchases can be represented by a point in space and the various points can be grouped into a system of indifference surfaces. On any one indifference surface lie all points which represent indifferent purchases. If one point lies on a higher indifference surface than another, the combination of purchases represented by the first point is preferable to that represented by the second.

In their normal form, the indifference surfaces are continuous and fit into each other as a system of non-intersecting shells. Further, the surfaces are downward sloping to each axis and convex to the origin at all points.

There are now two marginal rates of substitution at any point, corresponding to the substitution of X_3 for X_1 and for X_2 respectively. The first is the ratio of a small increase in X_3 to a small compensating decrease in X_1 (X_2 remaining fixed in amount), while the second is the ratio of the same increase in X_3 to a small compensating decrease in X_2 (X_1 remaining fixed in amount). The marginal rate of substitution of X_3 for X_1 is measured by the numerical gradient of the tangent plane to the indifference surface at any point taken in the direction perpendicular to Ox_2. The other marginal rate of substitution is similarly measured. The convexity of the indifference surfaces gives rise to the principle of increasing marginal rate of substitution in each case.

If the individual has a fixed sum to spend on the three goods, the market prices being also fixed, his purchases must be represented by some point on a plane which cuts the axes at A, B and C respectively. OA measures the amount of X_1 that can be purchased with the whole expenditure, OB the amount of X_2 and OC the amount of X_3 similarly. The numerical gradient of the plane perpendicular to Ox_2 is the ratio of the price of X_1 to the price of X_3, and the numerical gradient perpendicular to Ox_1 is the ratio of the price of X_2 to the price of X_3.

The fixed plane *ABC* cuts many indifference surfaces and touches one only at a point *E*. This point, as before, gives the individual's equilibrium purchases of the goods and is uniquely determined when the indifference surfaces are normal in form. Since the plane *ABC* is tangential to the indifference surface at *E*, the gradients of the plane are identical with the gradients of the tangent plane to the indifference surface at *E*. Hence, the marginal rate of substitution of X_3 for X_1 is equal to the ratio of the prices of X_1 and X_3 and the marginal rate of substitution of X_3 for X_2 is equal to the ratio of the prices of X_2 and X_3. The conditions given for the equilibrium purchases in the case of two goods thus readily extend to the case of three goods.

Finally, by taking a series of planes parallel to *ABC*, we obtain a curve in space on which all the equilibrium points, such as *E*, lie as total expenditure increases and prices remain constant. This curve shows the variation of the equilibrium purchases as total expenditure changes. It is usually the case that the curve slopes in the positive direction away from *O* and the expenditures on all three goods increase with the total expenditure. It is possible, however, that one, or even two, of the expenditures decreases as total expenditure increases. The form of the curve cannot be determined, in any case, without knowledge of the indifference surfaces.

In the completely general case of *n* goods, it is impossible to give a concrete picture of the preference scale of the individual or of the way in which his equilibrium distribution of expenditure is determined. We must, here, turn to a mathematical analysis of the problem and this analysis is given in the Mathematical Appendix below. It appears there that the equilibrium purchases of the individual are determined by conditions similar to those already given for two and for three goods, and that the distribution is stable if the preference scale is normal in form. The individual's distribution of expenditure follows from knowledge of the total expenditure and the market prices. It is also possible to trace the variation

of any one expenditure as total expenditure increases, prices being constant. Usually all expenditures increase with total expenditure, but one or more of them may decrease.

4. The Linear Preference Scale

One particular case of the preference scale of an individual consumer is of especial interest. In the foregoing analysis, no restrictions were imposed upon the nature of the preference scale, except those required to give the normal form of the indifference curves or surfaces. For particular purposes, we are at liberty to assume almost any definite form of the preference scale we think desirable. We now find it convenient to assume one very simple form and to work out the implications of this form in detail.

Taking the case of two goods, we assume that the marginal rate of substitution of X_2 for X_1, which describes the preference scale, is the ratio of two linear expressions in the amounts of the goods purchased. Algebraically, if R denotes the marginal rate of substitution and x_1 and x_2 the amounts of the goods purchased, we write

$$R = \frac{a_1 + a_{11}x_1 + a_{12}x_2}{a_2 + a_{21}x_1 + a_{22}x_2}$$

where the a's are constants. A preference scale defined by a marginal rate of substitution of this form can be termed a *linear preference scale*. Though limited in form, the linear preference scale can still vary to a considerable extent by allocating different values to the six constant a's.

For equilibrium, when the total expenditure is e and the market prices are p_1 and p_2, we have shown that the purchases of the individual must satisfy the relations

$$\frac{a_1 + a_{11}x_1 + a_{12}x_2}{a_2 + a_{21}x_1 + a_{22}x_2} = \frac{p_1}{p_2}$$

and $$p_1x_1 + p_2x_2 = e$$

Write
$$e_1 = p_1 x_1 \quad , \quad e_2 = p_2 x_2 \, ;$$
$$a_1 = p_1 b_1 \quad , \quad a_2 = p_2 b_2 \, ;$$
$$a_{11} = p_1^2 b_{11} \ , \ a_{12} = p_1 p_2 b_{12} \ , \ a_{21} = p_1 p_2 b_{21},$$
$$a_{22} = p_2^2 b_{22}.$$

The relations then appear
$$b_1 + b_{11} e_1 + b_{12} e_2 = b_2 + b_{21} e_1 + b_{22} e_2$$
and
$$e_1 + e_2 = e.$$
Eliminating $e_2 = e - e_1$, we have
$$b_1 + b_{11} e_1 + b_{12}(e - e_1) = b_2 + b_{21} e_1 + b_{22}(e - e_1)$$
i.e.
$$e_1 = k_1 e + c_1$$
where
$$k_1 = \frac{b_{22} - b_{12}}{b_{11} + b_{22} - b_{12} - b_{21}} \quad \text{and} \quad c_1 = \frac{b_2 - b_1}{b_{11} + b_{22} - b_{12} - b_{21}}$$
In the same way,
$$e_2 = k_2 e + c_2$$
where
$$k_2 = \frac{b_{11} - b_{21}}{b_{11} + b_{22} - b_{12} - b_{21}} \quad \text{and} \quad c_2 = \frac{b_1 - b_2}{b_{11} + b_{22} - b_{12} - b_{21}}$$

In the case of the linear preference scale, therefore, each of the expenditures depends on the total expenditure in a linear way. Further, by eliminating e from the two linear relations, we have
$$\frac{e_1}{k_1} - \frac{e_2}{k_2} = \frac{c_1}{k_1} - \frac{c_2}{k_2}$$
i.e.
$$\frac{p_1}{k_1} x_1 - \frac{p_2}{k_2} x_2 = \frac{c_1}{k_1} - \frac{c_2}{k_2}$$

This is the relation satisfied by the equilibrium purchases x_1 and x_2 for varying total expenditure, i.e. by the co-ordinates of points such as E_1, E_2, E_3, . . . of Figs. 3 and 4. The equation is linear and so the curve $E_1 \, E_2 \, E_3$ . . . shown in these figures reduces, in the case of a linear preference scale, to a straight line. The line slopes upwards to the right if both constants k_1 and k_2 are

positive, and downwards to the right or upwards to the left if one of the constants is negative.

The preference scale in the general case of n goods is defined to be linear if all the marginal rates of substitution are ratios of linear expressions in the purchases. In the Mathematical Appendix, it is shown that the expenditure on any one good is then linearly dependent on the total expenditure e according to the formula

$$e_r = k_r e + c_r \qquad (r = 1, 2, \ldots n)$$

where e_r is the expenditure on the rth good and where k_r and c_r are constants.

The constants k_r and c_r for each good are expressed in terms of co-efficients such as the b's defined above. It is to be remembered, however, that the b's depend, not only on the constants of the linear preference scale, but also on the fixed market prices. The constants thus apply only to a fixed preference scale and to a fixed set of prices and they vary in value as we pass from one set of tastes to another, or from one market to another.

For a complete set of goods entering into the individual's budget, the values of the k's and the values of the c's are each limited by one condition. Since the sum of the expenditures on the goods separately must equal the total expenditure, the relation

$$e = \sum_{r=1}^{n} e_r = \left(\sum_{r=1}^{n} k_r\right)e + \left(\sum_{r=1}^{n} c_r\right)$$

must hold whatever the value of e. Hence

$$\sum_{r=1}^{n} k_r = 1 \quad \text{and} \quad \sum_{r=1}^{n} c_r = 0.$$

The sum of a complete set of k's is equal to unity and of a complete set of c's zero. This is checked, in the case of two goods, by the above written expressions for the k's and c's. It follows that, if the k's are all positive, no one k can be greater than unity. Even if there are some negative k's, it is extremely unlikely that any one should be greater than, or indeed approach, unity. Further, some of the c's must be positive and some negative.

The dependence of the expenditure on any good upon the total expenditure can be represented on a diagram in which total expenditure is measured along the horizontal axis Oe and the expenditure on the rth good measured along the vertical axis Oe_r. In the case of the linear preference scale, the dependence appears as a straight line with gradient given by the constant k_r and cutting the vertical axis at a point which is a distance c_r above (or below) the origin. The lines for some goods (necessaries) cut the vertical axis above O and for other goods (luxuries) below O. The sum of the positive intercepts balances the sum of the negative intercepts.

The linear expenditure relation which corresponds to the linear preference scale is the basis of the work of the previous two chapters. In Chapter I, it is shown that the values of the constants k_r and c_r for any good can be related to average and marginal expenditures, to the concept of income elasticity of demand and to the definition of the order of urgency of the various goods. What has now been shown is the derivation of the simple expenditure formula from the postulates of the general theory.

It should be added that the special case of the linear preference scale can be of service even when the actual preference scale is not exactly linear in form. The actual scale may approximate more or less closely to the linear form over a limited region in which we happen to be interested. The results obtained by assuming the linear scale may then easily provide a sufficiently close description of the facts. The linear preference scale, it is here claimed, applies over a wider range of actual expenditure distributions than might appear at first sight, provided that approximate, rather than exact, results suffice and when only a fairly limited range of total expenditures is considered.

5. The Case of Independent Goods

Another particular case of the general preference scale can be considered briefly here, the case of what can be

called "independent goods." The concept of independent goods has occupied a curious, and rather ambiguous, position in the development of the theory of value. The early exponents of the theory, Jevons and Walras in particular, assumed independent goods throughout. It is only with Edgeworth and Pareto that the case of independent goods is reduced to an interesting special case of a more general theory in which goods are inter-related in consumption. According to Edgeworth and Pareto, a good is "independent" if its "marginal utility" depends only on the amount of the good purchased and not on the purchases of other goods. It is thus possible to have a single good as independent of all the others. The concept of "marginal utility" has, however, no essential place in the theory of value expounded here. Instead, we consider marginal rates of substitution (or the ratios of marginal utilities) as the main characteristic of a preference scale. The definition of independent goods given by Edgeworth and Pareto is, therefore, not appropriate and, if the case is to be retained, it must be defined afresh in terms of the significant characteristics of the preference scale.

A pair of goods X_1 and X_2 is defined as independent of other goods if the marginal rate of substitution between them is dependent on the purchases of the two goods but not on the purchases of other goods. This definition applies to any pair of goods whatever but no meaning is now attached to the independence of a single good.

The case of especial interest arises when all the pairs of the complete set of goods are independent in the sense just described, the case of a completely independent set of goods. In this case, it is shown in the Mathematical Appendix that the marginal rate of substitution between any two goods X_1 and X_2 is the ratio of an expression involving only the amount of X_1 purchased to an expression involving only the amount of X_2 purchased. In fact, the preference scale of an individual, for whom the goods form an independent set, is described by a set of functional expressions each involving only one of the

amounts purchased. The conditions which determine the equilibrium purchases are now rather more easily managed in algebraic terms. It can be shown, however, that the results in the sphere of the observed demands of the individual are in no essential way simpler in the independent case than in the full case of inter-related goods.[1]

When the goods form an independent set and we also take the special linear case of the preference scale, we obtain a case that can be handled algebraically with great ease. The co-efficients such as b_{12} and b_{21} that appear in the conditions for equilibrium in the case of the linear preference scale now disappear. The constants k_r and c_r of the linear expenditure relation for the rth good are given by rather simpler expressions. The complete development is provided in the Mathematical Appendix, where it appears that k_r depends only on the co-efficients such as b_{11} and b_{22}, while c_r also involves the co-efficients such as b_1 and b_2.

Suppose, now, that the values of the k's and the c's for all goods have been determined from statistical data of family budgets. In the full case of the linear preference scale, it is not possible to deduce the values of the b's which involve the co-efficients of the preference scale (as well as the known market prices). But, in the case where the goods are also independent, the number of the b's is smaller and knowledge of the k's and c's enables us to find all the b's when values are allotted to two of them, say b_1 and b_{11}. The main characteristics of the actual preference scale, in the linear and independent case, can thus be evaluated from budget data.

The practical usefulness of this conclusion is not very great. It applies only when the preference scale is known to be independent as well as linear. But there is

[1] See Hicks and Allen, *op. cit.*, pp. 74–6 and 214–18. The case of independent goods is not, as has so long been thought, the most interesting case in the theory of value; it is comparatively unimportant. A certain simplification is to be noticed, however, in the rather artificial case where only two goods appear in the individual's budget and the goods are independent.

no way of testing the applicability of the assumption of independence; all the information obtainable from the data has been used to test the hypothesis of linearity and to give the values of the k's and c's, and there can be no relations left to tell us whether the goods can be taken as independent or not. All that we can do is to assume independence on *a priori* grounds, a step we are scarcely prepared to take. The case of independent goods would appear to be of little practical interest.

6. The Variation of Tastes

We have given a theoretical analysis of the distribution of expenditure and of the changes in the distribution as total expenditure changes. This analysis provides a useful, if not indispensable, basis for the statistical investigation of actual budget data. Such data relate to the budgets of individual families with varying total expenditure but subject to what can be taken as fixed market conditions and, in particular, subject to fixed market prices. From the data, therefore, we obtain information concerning the way in which the expenditure distribution of a family changes as the level of total expenditure changes. If the data are given, as in the cases considered in Chapter I, only in average form, the information obtainable is then necessarily limited. On the other hand, if the complete data for all families are available, as in Chapter II, then much more information can be expected.

In relating the concepts of the theoretical analysis to the actual statistical data, we must consider one very important fact. The theoretical analysis describes the expenditure distribution of a given family whose purchases can be considered as based on a fixed preference scale. The analysis is thus only of direct application to budget data collected from a group of families known to have the same complex of preferences. In practice, however, the needs and tastes of the families investigated vary to such an extent that the assumption of a common preference scale is clearly inappropriate. A considera-

tion of the theoretical aspects of varying needs and preferences is necessary before we can proceed.

The complex of preferences of families included in any budget collection varies, not only because of variation in the inherent tastes of the individuals in the family, but also because of the variation in the age and sex composition of the family. From the point of view of the present study at least, the variation of the family preference scale due to the variation in family composition is almost entirely irrelevant. Methods must, therefore, be devised to eliminate the effect of this variation from the budget data and the practical devices found most suitable for this purpose are sufficiently described in Chapter I, § 4, above. We can now proceed on the assumption that the budget data have been corrected for varying family composition, either by the inclusion of an index of family composition in all our relations (so that we can consider a family of any given composition), or by the relatively simple process of taking all expenditures per equivalent adult. In this way, we isolate, as far as possible, the effect of the variation of inherent tastes in the expenditures of the group of families investigated.

It may happen that, after the elimination of the effect of family composition, all families are found to make their purchases according to the same scale of preferences. This case is very unlikely to arise in practice, even as a first approximation, but the application of the theory to the simple case must be given first as a basis for a later application to more realistic cases.

Theory now suggests that the expenditure of any family on a selected item depends in some definite way upon the total expenditure of the family. If this dependence is not shown, within the limits set by observation errors, by the data of the actual budget collection, then the theory is not applicable to the group of families concerned in the sense that their purchases cannot be made according to any definite preference scale. If the dependence is shown by the data, then its form can be investigated over the limited expenditure range covered

by the families. In particular, it can be determined whether the expenditure relation is linear, i.e. whether the preference scale common to all families takes the special linear form. Alternatively it can be determined whether the linear expenditure relation, and the linear preference scale, is approximately appropriate to any part of the income range. The data can be tested in this way whether given in average or in complete form.

If the common preference scale is found to be at least approximately linear, the variation of any expenditure is described by the relation

$$e_r = k_r e + c_r$$

and the values of the constants can be determined at once from the data. Expressions for the derived measurements of average expenditure w_r and elasticity of expenditure η_r are then obtained. These expressions are dependent on e and only take definite values when the expenditure level is specified. In summarising data to which the linear relation is found to apply, it is convenient to give the values of w_r and η_r at the average level of total expenditure $\bar{e}$ found in the group of families. These values, $\bar{w}_r$ and $\bar{\eta}_r$, are dependent on the choice of families making up the budget collection, a fact which has already been noted in Chapter I, § 2, above.

As we pass from one group of families with a common preference scale to another group with a different preference scale (e.g. from working-class to middle-class families), different values of the characteristics k_r, c_r, $\bar{w}_r$ and $\bar{\eta}_r$ are obtained. The same applies when we pass from one market to another even if the preference scale remains unaltered. In fact, the derivation of different values of these concepts from two groups of families indicates that the families make their purchases either according to different preference scales or on markets with different relative prices.

In general, however, even with families of fairly homogeneous composition as regards (e.g.) social position, it is inevitable that inherent tastes vary from family to family.

The problem is to discover how far our theoretical concepts are applicable to the results of budget data when tastes vary.

If the data are presented only in average form, we have very little to guide us. The results of the statistical investigation can only be related to the theoretical concepts by making the highly abstract assumption that the various families can be represented adequately, as far as the distribution of expenditure is concerned, by a definite though fictitious "average" family. To this "average" family, there corresponds an "average" preference scale and an "average" relation between the expenditure on any item and total expenditure. Our assumption, then, is that the complexity and variety of tastes amongst the families is limited so that the average expenditures on an item for different groups of total expenditure conform to an average expenditure relation. The most we can derive from average data, therefore, is an indication that average expenditures do or do not conform to definite relations. If the data are given in a sufficient number of expenditure groups and if a definite set of relations is shown, then we have some support for the theoretical construction of preference scales interpreted in the average sense indicated. Further, if the average expenditure relations are found to be linear, we can say that the distribution of expenditure on the average conforms with the assumption of a linear preference scale for an average family. The analysis of budget data on these lines is given in Chapter I. The results there obtained are, however, useful in the description of the variation of expenditure even if the applicability of the theoretical construction is not certain.

It is possible to proceed further with the analysis of the variation of tastes when the budget data give complete information of the expenditures of all families. Suppose, now, that the preference scale of each family is known to be linear in form but that the scale varies from family to family. There is, therefore, a linear relation between any item of expenditure and the total

expenditure for each family but the relation is different for different families. If the expenditures e_r on the item X_r are plotted, for the various families, against the corresponding total expenditures e as points on a scatter diagram, each point lies on an expenditure line which varies in position from point to point (see Fig. 5).

The variation of the position of expenditure lines is derived from the variation of the co-efficients $b_1, b_2, \ldots,$ $b_{11}, b_{22}, b_{12}, \ldots$ obtained in the case of the linear preference scale (§ 4, above). As a first case, suppose that

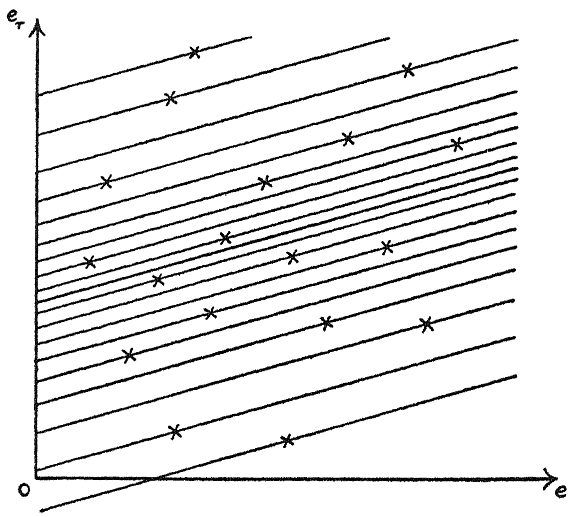


FIG. 5.

only the co-efficients $b_1, b_2, \ldots$ vary while the other co-efficients are the same for all families. It is shown in the Mathematical Appendix that the expenditure lines for different families are then parallel, i.e. we have

$$e_r = k_r e + c_r$$

where k_r is the same but c_r varies from one family to another. The situation is illustrated by Fig. 5. The distribution of c_r about its average value is normal if the distribution of b_1, of $b_2, \ldots$ are all normal over the range of families, or if the number of items is large and

the variations of b_1, b_2, . . . are independent, or if both these conditions hold. Further, if there is no correlation between total expenditure and the variation of b_1, b_2, . . . and if the number of families is large, the distribution of values of c_r is normal *in any total expenditure group* and its standard deviation is the same for all expenditure groups, i.e. the distribution of c_r over all families is "homoscedastic."

In the more general case, the co-efficients b_{11}, b_{22}, b_{12}, . . . vary over the families as well as the co-efficients b_1, b_2, . . . The expenditure lines of the different families in the scatter diagram are then no longer parallel and the variation of tastes is more complicated. If, however, we can assume that the variation of the co-efficients b_1, b_2, . . . is greater than the variation of the other co-efficients, then the expenditure lines will not vary greatly in slope as compared with their variation in height.[1]

In such cases, therefore, it is appropriate to fit an average expenditure line ($e_r = k_r e + c_r$) to the scatter diagram, by least squares or otherwise, and to represent the actual expenditure of any family by the relation

$$e_r = k_r e + c_r + v_r$$

where v_r, which takes positive and negative values about a zero mean, is the difference between the actual expenditure and the average expenditure at the same level of total expenditure. In the case where the expenditure lines are exactly parallel, we have seen that the distribution of v_r, which is the same as the distribution of the previous c_r about its mean, is normal and homoscedastic under the conditions stated. In the more general case where the lines differ little in slope, the distribution of c_r is approximately normal and homoscedastic.

[1] There is some reason to expect the co-efficients b_1, b_2, . . . to be large in comparison with the others in linear preference scales found in practice. The former fix the average value of the marginal rates of substitution while the latter are necessary to describe the variation of the rates for different consumption levels. If the curvature of the indifference curves or surfaces is not great, the marginal rates of substitution change slowly and our assumption is justified.

In Chapter II, we investigated scatter diagrams of the kind now considered and it was found that the distribution of the residual expenditures, representing the variation of tastes, could be described by the normal curve of error to a fair approximation in most cases. There was also some indication that the distribution could be taken as homoscedastic, i.e. that the variation of v_r within any one group of total expenditure has a constant standard deviation. To establish this latter point, however, we require budget collections covering many more families, and more reliable in nature, than any to which we have had access. We have now set out one set of theoretical conditions, relating to the variation of the linear preference scale from family to family, which accounts for such a variation of residuals. There may be, of course, other sets of conditions producing the same result.[1]

If the observed normal and (possibly) homoscedastic distribution of v_r is due to the applicability of the conditions enumerated, we have succeeded in giving an account of the variation of tastes to a large extent. Further, this is also the justification we have for the use of the average expenditure relation in the description of groups of families. The constant k_r of this relation is almost the same for all families and only the constant c_r varies, in a known way about its average value obtained from the average relation, from family to family.[2] On the other hand, if the observed distribution of residuals is not normal, and particularly if it is not symmetrical, then it is probable that there are some biased factors in the variation of tastes of which we have taken no account.

[1] Almost any set of small, sporadic and independent variations in tastes, if sufficiently numerous, will give rise to an approximately normal distribution of residual expenditure. It would appear, however, that the second property of homoscedasticity is only obtained under more special conditions. The conditions set out above account for the homoscedastic, as well as for the normal, property of the distribution of residuals.

[2] It should be noticed that the variation of the constants c_r ($r = 1, 2, \ldots n$) for different families implies that the order of urgency of the budget items may vary from family to family. But there is a definite "average" order of urgency for the group of families, i.e. the order defined by the constants of the average expenditure relations.

In any case, we have here to look elsewhere for a description of the variation of tastes.

In conclusion, it should be noticed again that the co-efficients denoted by the b's depend on the market prices as well as on the characteristics of the preference scale of the family concerned. An exactly similar situation to that discussed above is thus obtained when market prices vary from family to family as well as inherent tastes. In practice, the market to which the budgets relate is always of some geographical extension and prices are never completely uniform. In any case, there is an obvious connection between tastes and the prices and qualities of commodities bought. Our analysis of the variability of tastes, as reflected in expenditures, covers all cases.

7. Summary of the Applications of the Theory

The object of our statistical enquiry into family budget data, in so far as it is more than merely descriptive, is to test whether the theoretical analysis based on preference scales is applicable to the distributions of expenditures actually found. If the theory is applicable, the data can then provide a description of the actual form assumed by the relation between the expenditure on one item and the total expenditure that theory suggests. In particular the data can be used to test whether the particular linear form of the preference scale and expenditure relation is appropriate to the families investigated, at least approximately and over the range of incomes shown. Finally the main characteristics of the expenditure relations found in the budget data can be evaluated and comparisons can be made between budget data collected at different dates, in different countries and under different market conditions or for different social classes.

If the data are provided in average form, we can only test the applicability of the theory with the additional assumption of an average individual with an average preference scale and expenditure relation. It can be seen whether average expenditures on various items bear

linear or other relations to total expenditure. If such relations are shown, we have some support for the theory and we can proceed fairly safely to the evaluation of the characteristics and concepts which the theory suggests are important. If the relations are not shown, we can only assume that there is no average appropriate to the family group or that the whole theoretical structure is of doubtful applicability. It has been seen, in Chapter I, that there is a good fit, allowance being made for observation and sampling errors and for imperfections in the data, to a linear expenditure relation and occasionally to a parabolic relation. In the derivation of the values of the constant marginal expenditure k and of the income elasticity of demand, we are then able to make Engel's law precise and to classify goods according to a scale of urgency. We are also able to find certain special items of expenditure, such as rye bread, flour and margarine, where expenditure decreases absolutely as the total expenditure increases. This is a possibility envisaged in the theoretical analysis and it is important that its practical existence should be established.

In the cases where the budget data are given in complete and detailed form, we are in a position to subject the theory to a more searching test, both as regards its applicability and in respect of its power to account for the facts. Our main object here is to obtain some description of the variation of preferences over the families included in the group investigated. The data examined in Chapter II shows that there is considerable variation of expenditure about the average for the group but that the variation, when the effect of variation of needs and income has been allowed for, can be described, at least approximately, by the normal law. There is at least one set of theoretical conditions, based on the postulate of a preference scale variable from family to family, which accounts both for the appropriateness of an average expenditure relation in the description of the group and for the normality of the variation about the average. Our attempt to reduce the average expenditures to simple

laws and to account for individual variations from the average has not been entirely unsuccessful.

Finally, what we have been able to describe statistically has been the expenditure relation and its characteristics. The preference scale of the individual, on which his purchases are assumed to be based, remains hidden from us. Its characteristics are not amenable to statistical evaluation, at least by means of budget material. It has been shown, in fact, that it is conceivable that the features of the preference scale can be measured in one case only, the case where we assume that all goods are independent in consumption. But we neither expect this case to hold nor have we any means of determining whether it does hold in any practical case.

8. Price Elasticities of Demand

Though the primary object of investigations into budget data must be the description of changes in the distribution of family expenditure as income changes, yet the results of these investigations are not without application to other problems. There is, notably, the problem of the variation of demand due to changes in the price structure, a variation described by the familiar price elasticities of demand. It is shown in the Mathematical Appendix that a price elasticity of demand depends on two factors, the first being an "expenditure effect" arising from the change in real income which accompanies any price change and the second being a "substitution effect" arising from the change in relative prices. The analysis of budget data provides some information concerning the first of these effects.

If the preference scale of an individual is linear, it is shown that the expenditure effect in the individual's (direct) price elasticity of demand for any good X_r is measured by the constant k_r found in the linear expenditure relation. Certain conclusions can be drawn from this fact since the value of the constant is less than unity and usually quite small.

If there is no possibility of substituting other goods for

X_r in consumption, then the price elasticity of demand for X_r is equal to k_r, i.e. the elasticity is less than unity and usually quite small. The demand for such a good is inelastic in response to changes in its own price. It follows that elastic demands, or even moderately inelastic demands, can only arise in the case of goods for which considerable substitution is possible. For most items in the individual's budget, and particularly for most groups of items such as meat, dairy produce or clothing (where substitution effects are to be expected to be small), the demand is certain to be inelastic with respect to changes in the price of the item.[1]

The price elasticity of demand is, however, dependent on the level of income or total expenditure, amongst other factors. But, in our linear case, the first term k_r is unaffected by income changes and the price elasticity of demand is only modified by changes in the substitution factor as income changes. It is to be expected, moreover, that substitution becomes more easy for most goods as income rises. The larger expenditure is spread over a wider range of items and the possibilities of substituting other items for a given item are thereby increased. It follows that the elasticity of demand for any item with respect to changes in its price is likely to increase with income. Demands tend to become more elastic as the income level rises.[2]

Further, it has been shown that the value of k_r is negative for certain particular items of expenditure, i.e. that expenditure on an item can decrease as income rises. In this case, the expenditure effect in the price elasticity of demand for the item is negative and it may happen that the price elasticity is itself negative. This possi-

[1] The term "inelastic" is used when the measurement of elasticity is less than unity; and the term "elastic" when the measurement is greater than unity.

[2] If this point is correct and of wide applicability, it has a bearing upon the usual statistical devices adopted for the evaluation of price elasticities. It is assumed by Professors Pigou and Frisch, Dr. Leontief and others that a curve showing demand against the price of the item shifts with unaltered elasticity over time as the real incomes of the consumers (amongst other factors) change. This assumption is one of doubtful validity.

bility is greater the larger the numerical value of the negative k and the smaller the substitution effect between other goods and the one concerned. The possibility of demand for an item falling in response to a decrease in the price is thus established.

It is also shown in the Mathematical Appendix that the value of the constant c_r, which appears in the linear expenditure formula, can be written as dependent on the elasticities of substitution between the good X_r and other goods. From this fact, it can be deduced that X_r can only be a luxury (c_r negative) if it is in competition with many of the other items of the individual's budget. On the other hand, X_r is a necessary (c_r positive) if there are considerable complementary relationships between the good and other goods. In fact, as we should expect, the scale of urgency of items in the budget is closely connected with the competitive and complementary relations of the items in consumption.

MATHEMATICAL APPENDIX

1. The Preference Scale of an Individual Consumer.

The following notation is adopted. The total expenditure of an individual consumer is e. There are n consumers' goods X_1, X_2, . . . X_n obtainable on a given market at uniform prices $p_1, p_2, \ldots p_n$ respectively. The purchases of the individual are $x_1, x_2, \ldots x_n$ of the n goods respectively and his expenditures on the goods are $e_1, e_2, \ldots e_n$. Hence

$$e_r = p_r x_r \quad \text{and} \quad \sum_{r=1}^{n} e_r = e$$

The fundamental assumption is that the relations of the goods in the individual's consumption are expressed by means of a definite scale of preferences which serves to distinguish combinations of purchases $(x_1, x_2, \ldots x_n)$ according to preference or indifference. In particular, each possible combination of purchases can be put into one or other of a series of sets of indifferent combinations. The different sets can be arranged in ascending order of preference so that one level of preference is associated with each set and the level increases as we proceed through the order of the sets.

A numerical index is now attached to each indifferent set in such a way that the index increases with the level of preference. The index is some function of $x_1, x_2, \ldots x_n$ which takes a constant value for all indifferent combinations of purchases and increases in value as the level of preference increases. The association of the numerical index with the preference levels is designed to preserve the *order* of preference but does not attempt to *measure* preference. The index, in whatever form it takes, can thus be represented by a function

$$u = F\{\phi(x_1, x_2, \ldots x_n)\} \quad . \quad . \quad . \quad (1)$$

where $\phi(x_1, x_2, \ldots x_n)$ is any one form of the index and where F is an arbitrary function limited simply by the fact that $F'(\phi) > 0$.

By definition, if u_1 and u_2 are the values of the index for two combinations of purchases, then

$$u_1 = u_2$$

if the combinations are indifferent and at the same preference level, and

$$u_1 > u_2$$

if the first combination is in a higher indifferent set, at a higher level of preference, than the second.

The function u is called the *function index of utility*. A set of indifferent combinations of purchases is defined by the relation

$$\phi(x_1, x_2, \ldots x_n) = \text{constant}$$

where the value of the constant simply determines the particular level of preference concerned. This relation is represented diagrammatically by a system of *indifference loci* in n-dimensional space, i.e. by indifference curves (two goods), indifference surfaces (three goods) and so on.

As the combination of purchases varies in any way, the variation in the value of u is given by the partial derivatives

$$\frac{\partial u}{\partial x_1} = F'(\phi)\frac{\partial \phi}{\partial x_1}; \ \frac{\partial u}{\partial x_2} = F'(\phi)\frac{\partial \phi}{\partial x_2}; \ldots ; \frac{\partial u}{\partial x_n} = F'(\phi)\frac{\partial \phi}{\partial x_n}.$$

The arbitrary element represented by the function F appears in these expressions. But it is to be noticed that the ratios

$$\frac{\partial u}{\partial x_1} : \frac{\partial u}{\partial x_2} : \ldots : \frac{\partial u}{\partial x_n}$$

are independent of F and can be written as

$$\frac{\partial \phi}{\partial x_1} : \frac{\partial \phi}{\partial x_2} : \ldots : \frac{\partial \phi}{\partial x_n}$$

obtained from any one form of the index $\phi(x_1, x_2, \ldots x_n)$.

The partial derivatives of the function u can be called the *marginal utilities* of the goods $X_1, X_2, \ldots X_n$ respectively. The marginal utilities themselves are of little use, however, since the arbitrary function F is necessarily involved in their expression. It is the ratios of the marginal utilities that are of importance. We replace the marginal utility concepts, therefore, by the more definite concepts

$$R_1 = \frac{\dfrac{\partial \phi}{\partial x_1}}{\dfrac{\partial \phi}{\partial x_n}}; \ R_2 = \frac{\dfrac{\partial \phi}{\partial x_2}}{\dfrac{\partial \phi}{\partial x_n}}; \ldots ; R_{n-1} = \frac{\dfrac{\partial \phi}{\partial x_{n-1}}}{\dfrac{\partial \phi}{\partial x_n}}$$

Each of these $n - 1$ ratios is a function of the purchases $x_1, x_2, \ldots x_n$ which can be interpreted in the following way.

For any variation in the purchases of the two goods X_1 and X_n only, we have

$$du = F'(\phi)\left(\frac{\partial\phi}{\partial x_1}dx_1 + \frac{\partial\phi}{\partial x_n}dx_n\right) = F'(\phi)\frac{\partial\phi}{\partial x_n}(R_1 dx_1 + dx_n)$$

If the variation leaves the preference level unaltered, $du = 0$ and

$$R_1 dx_1 + dx_n = 0$$

i.e.
$$R_1 = -\,dx_n/dx_1$$

The value of the function R_1 at any consumption level is thus equal to the increment in X_n necessary to compensate for a given unit decrement in X_1, the level of preference remaining unaltered. The function R_1 is, therefore, called the *marginal rate of substitution of X_n for X_1*. Its values describe an aspect of the preference scale which corresponds to one level of indifference. The other functions, $R_2, R_3, \ldots R_{n-1}$, are to be interpreted in a similar way.

The individual's preference scale is thus expressed by a set of functions $R_1, R_2, \ldots R_{n-1}$ which give the marginal rates of substitution of the good X_n for the other goods $X_1, X_2, \ldots X_{n-1}$ at any level of consumption. The forms of the functions are independent of any arbitrary element and determine between them the form of the complete preference scale. The further assumption can now be added that $R_1, R_2, \ldots R_{n-1}$ is each a continuous function of continuous variables $x_1, x_2, \ldots x_n$. This is an assumption adopted purely for mathematical convenience. The possibility of discontinuities in the preference scale remains; it is assumed away, for convenience, in the following mathematical analysis.

In the case of two goods X_1 and X_2, the utility function index is

$$u = F\{\phi(x_1, x_2)\}$$

and the equation of the system of indifference curves is

$$\phi(x_1, x_2) = \text{constant.}$$

At any point (x_1, x_2), the tangent line to the indifference curve has a gradient (referred to Ox_1) given by

$$\frac{dx_2}{dx_1} = -\frac{\dfrac{\partial\phi}{\partial x_1}}{\dfrac{\partial\phi}{\partial x_2}} = -R$$

where R is the marginal rate of substitution of X_2 for X_1. The values of the marginal rate of substitution between the goods are

thus represented, in diagrammatic terms, by the numerical gradients of the indifference curves at various points. Similarly, in the case of three goods, the tangent plane to the indifference surface of the system

$$\phi(x_1, x_2, x_3) = \text{constant}$$

at any point is given by the two gradients

$$\frac{dx_3}{dx_1} = -\frac{\dfrac{\partial\phi}{\partial x_1}}{\dfrac{\partial\phi}{\partial x_3}} = -R_1$$

$$\frac{dx_3}{dx_2} = -\frac{\dfrac{\partial\phi}{\partial x_2}}{\dfrac{\partial\phi}{\partial x_3}} = -R_2$$

where 3 takes the place of n written in the general case above.

The two marginal rates of substitution at any consumption level are thus represented by the numerical values of the two basic gradients of the indifference surface through the point in question.

2. The Conditions for Equilibrium

The problem is to determine the equilibrium purchases of the individual when his total expenditure e and the market prices $p_1, p_2, \ldots p_n$ are all given. It is taken that these purchases are such that the individual attains the highest possible position on his preference scale, i.e. such that u has a maximum value, consistent with the given total expenditure and the given prices. The values of $x_1, x_2, \ldots x_n$ are to be determined, therefore, so that

$$du = 0 \quad \text{and} \quad d^2u < 0$$

subject to

$$p_1x_1 + p_2x_2 + \ldots + p_nx_n = e.$$

The necessary condition ($du = 0$) is

$$F'(\phi)\frac{\partial\phi}{\partial x_n}(R_1dx_1 + R_2dx_2 + \ldots + dx_n) = 0$$

i.e. $$R_1dx_1 + R_2dx_2 + \ldots + dx_n = 0$$

since $F'(\phi)$ is positive (not zero), and it is assumed that $\dfrac{\partial\phi}{\partial x_n}$ is not zero.

This equation is subject to the condition that

$$p_1 dx_1 + p_2 dx_2 + \ldots + p_n dx_n = 0$$

The conditions necessary for equilibrium are thus

$$\frac{R_1}{p_1} = \frac{R_2}{p_2} = \ldots = \frac{R_{n-1}}{p_{n-1}} = \frac{1}{p_n}$$

$$\left.\begin{array}{l} \text{i.e.} \quad R_1 = p_1/p_n;\ R_2 = p_2/p_n;\ \ldots;\ R_{n-1} = p_{n-1}/p_n \\ \text{together with} \\ \qquad p_1 x_1 + p_2 x_2 + \ldots + p_n x_n = e \end{array}\right\} \quad (2)$$

Equilibrium is only possible when the purchases are such that the marginal rates of substitution are brought into equality with the corresponding given price ratios.

A position of the kind satisfying the conditions (2) is one of maximum preference, as opposed to (e.g.) minimum preference, only if the condition $d^2u < 0$ is also satisfied subject to the same limiting condition as before. This restriction must now be examined. We have

$$\begin{aligned} d^2u &= d\left\{F'(\phi)\frac{\partial\phi}{\partial x_n}(R_1 dx_1 + R_2 dx_2 + \ldots + dx_n)\right\} \\ &= d\left\{F'(\phi)\frac{\partial\phi}{dx_n}\right\}.(R_1 dx_1 + R_2 dx_2 + \ldots + dx_n) \\ &\quad + F'(\phi)\frac{\partial\phi}{\partial x_n} d(R_1 dx_1 + R_2 dx_2 + \ldots + dx_n) \end{aligned}$$

But we already have

$$R_1 dx_1 + R_2 dx_2 + \ldots + dx_n = 0$$

Hence, d^2u is proportional to $d(R_1 dx_1 + R_2 dx_2 + \ldots + dx_n)$ and the limiting condition for a maximum is that

$$d(R_1 dx_1 + R_2 dx_2 + \ldots + dx_n) < 0$$

i.e. that

$$\begin{aligned} &\frac{\partial R_1}{\partial x_1} dx_1^2 + \frac{\partial R_1}{\partial x_2} dx_1 dx_2 + \ldots + \frac{\partial R_1}{\partial x_n} dx_1 dx_n \\ &+ \frac{\partial R_2}{\partial x_1} dx_1 dx_2 + \frac{\partial R_2}{\partial x_2} dx_2^2 + \ldots + \frac{\partial R_2}{\partial x_n} dx_2 dx_n \\ &\qquad \cdots\cdots \\ &+ \frac{\partial R_{n-1}}{\partial x_1} dx_1 dx_{n-1} + \frac{\partial R_{n-1}}{\partial x_2} dx_2 dx_{n-1} + \ldots + \frac{\partial R_{n-1}}{\partial x_n} dx_{n-1} dx_n \end{aligned}$$

is a negative definite quadratic form subject to

$$R_1 dx_1 + R_2 dx_2 + \ldots + dx_n = 0$$

It is known that the condition is equivalent to the fact that the determinant

$$\begin{vmatrix} R_1 & R_2 & \dots & R_{n-1} & 1 \\ \dfrac{\partial R_1}{\partial x_1} & \dfrac{\partial R_1}{\partial x_2} & \dots & \dfrac{\partial R_1}{\partial x_{n-1}} & \dfrac{\partial R_1}{\partial x_n} \\ & & \dots\dots & & \\ \dfrac{\partial R_{n-1}}{\partial x_1} & \dfrac{\partial R_{n-1}}{\partial x_2} & \dots & \dfrac{\partial R_{n-1}}{\partial x_{n-1}} & \dfrac{\partial R_{n-1}}{\partial x_n} \end{vmatrix}$$

must have principal minors of successive orders which are alternatively positive and negative.[1] The determinant can be written

$$\frac{R_1 R_2 \dots R_{n-1}}{x_1 x_2 \dots x_n} \begin{vmatrix} x_1 R_1 & x_2 R_2 & \dots & x_{n-1} R_{n-1} & x_n \\ \rho_{11} & \rho_{12} & \dots & \rho_{1n-1} & \rho_{1n} \\ & & \dots\dots & & \\ \rho_{n-11} & \rho_{n-12} & \dots & \rho_{n-1n-1} & \rho_{n-1n} \end{vmatrix}$$

where $$\rho_{rs} = \frac{x_s}{R_r} \frac{\partial R_r}{\partial x_s}$$

$$r = 1, 2, \dots (n-1) \text{ and } s = 1, 2, \dots n$$

The conditions for a maximum are thus

$$R_1 > 0; \quad \begin{vmatrix} x_1 R_1 & x_2 R_2 \\ \rho_{11} & \rho_{12} \end{vmatrix} < 0; \quad \begin{vmatrix} x_1 R_1 & x_2 R_2 & x_3 R_3 \\ \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{21} & \rho_{22} & \rho_{23} \end{vmatrix} > 0; \text{ etc.} \tag{3}$$

These can be called *stability conditions*; if they are satisfied, the equilibrium position defined by the conditions (2) is stable in the sense that it is one of maximum preference.

The stability conditions (3) relate to the form of the preference scale at the equilibrium position. The conditions may be satisfied by some preference scales and not by others. The "normal" preference scale is now defined as any scale which satisfies the stability conditions (3) at all positions on the scale. It follows, in this case of a normal preference scale, that a unique and stable equilibrium set of purchases is defined whatever the market prices or the total expenditure of the individual may be. For other preference scales, multiple equilibrium positions are possible and some of the positions will be stable and some unstable.

[1] See Bocher, *Higher Algebra*, Chaps. X and XI.

In the case of two goods, the stability conditions (3) are

$$R > 0 \quad \text{and} \quad \begin{vmatrix} x_1 R & x_2 \\ \rho_{11} & \rho_{12} \end{vmatrix} < 0$$

where R is the marginal rate of substitution of X_2 for X_1. From the indifference curve system defined by the relation

$$\phi(x_1, x_2) = \text{constant}$$

we obtain

$$\frac{dx_2}{dx_1} = -R$$

and

$$\frac{d^2x_2}{dx_1^2} = -\frac{dR}{dx_1} = -\left\{\frac{\partial R}{\partial x_1} + \frac{\partial R}{\partial x_2}\frac{dx_2}{dx_1}\right\}$$

$$= R\frac{\partial R}{\partial x_2} - \frac{\partial R}{\partial x_1}$$

$$= \frac{R}{x_1 x_2}\left\{Rx_1\left(\frac{x_2}{R}\frac{\partial R}{\partial x_2}\right) - x_2\left(\frac{x_1}{R}\frac{\partial R}{\partial x_1}\right)\right\}$$

$$= \frac{R}{x_1 x_2}\begin{vmatrix} x_1 R & x_2 \\ \rho_{11} & \rho_{12} \end{vmatrix}$$

The stability conditions are thus

$$\frac{dx_2}{dx_1} < 0 \quad \text{and} \quad \frac{d^2x_2}{dx_1^2} < 0$$

In diagrammatic terms, the indifference curve through any point of the Ox_1x_2 plane must be downward sloping and convex to the origin. This is the normal form of the indifference curve system. It follows that, as we move along an indifference curve substituting X_2 for X_1, the numerical gradient of the curve (referred to Ox_1) increases. In other words, as X_2 is substituted for X_1, the marginal rate of substitution of X_2 for X_1 increases. The stability conditions, therefore, include what is known as the principle of increasing marginal rate of substitution.

In the case of three goods, it follows, in an analogous way, that the indifference surface through any point of the $Ox_1x_2x_3$ space is downward sloping and convex to the origin. This stability condition involves the principle of increasing marginal rate of substitution for both the substitutions (X_3 for X_1 and X_3 for X_2) considered in this case.

3. The Expenditure Function and its Variation

The equilibrium conditions (2) are n in number and sufficient for the determination of the n unknown purchases $x_1, x_2, \ldots x_n$ of the individual when the values of $e, p_1, p_2, \ldots p_n$ are given. For different values of these parameters, the equilibrium conditions give different equilibrium purchases. The purchases of the n goods made by the individual in equilibrium are each given as a function of the variable parameters $e, p_1, p_2, \ldots p_n$ and we can write

$$x_r = F_r(e, p_1, p_2, \ldots p_n) \qquad (r = 1, 2, \ldots n)$$

These are the individual's demand functions. In the case of the normal preference scale, the stability conditions (3) are always satisfied and a unique and stable equilibrium set of purchases results whatever the values of $e, p_1, p_2, \ldots p_n$. In this case, the demand functions are single-valued.

The demand functions represent the result of comparing different equilibrium positions corresponding to different total expenditures and to different market prices. The comparison is now limited to the case where the individual has different expenditures but where the market prices are known and fixed, the case appropriate to studies of budget data. We have

$$e_r = p_r x_r = p_r F_r(e, p_1, p_2, \ldots p_n)$$

and, since the prices are now fixed, we can write

$$e_r = f_r(e) \qquad (r = 1, 2, \ldots n)$$

These are the individual's *expenditure functions*. The form of each function f_r depends on two factors—firstly, the form of the individual's preference scale and, secondly, the set of market prices that happen to be fixed for him.

It will usually be the case that the demand for, and the expenditure on, any good X_r increases as the total expenditure increases. Here, e_r is an increasing function of e and the derivative de_r/de is positive. But, in exceptional cases, it can happen that the expenditure on X_r decreases as total expenditure increases; e_r is then a decreasing function of e and the derivative is negative.

From the expenditure function for any good X_r, we derive the following useful concepts:

Marginal expenditure on X_r $\quad k_r = \dfrac{de_r}{de} \qquad \left(\sum_{r=1}^{n} k_r = 1\right)$

Average expenditure on X_r $\quad w_r = \dfrac{e_r}{e} \qquad \left(\sum_{r=1}^{n} w_r = 1\right)$

Elasticity of expenditure on X_r $\quad \eta_r = \dfrac{ede_r}{e_r de} = \dfrac{k_r}{w_r}$

$$\left(\sum_{r=1}^{n} w_r \eta_r = 1\right)$$

Each of these derived concepts is dependent on the level of total expenditure e. They are all expressed in ratio form and are quite independent of the monetary units chosen for measuring prices and expenditures. Between them, the concepts serve to describe the variation of the equilibrium distribution of expenditure over the different goods as the individual's total expenditure varies, prices remaining constant.

It should be noticed that the elasticity of expenditure η_r is the elasticity of the demand function x_r with respect to changes in the variable e. In fact, since the prices are constant,

$$\eta_r = \frac{ede_r}{e_r de} = \frac{e}{p_r x_r} \frac{d(p_r x_r)}{de} = \frac{e}{x_r} \frac{dx_r}{de}$$

Our η_r is thus the income elasticity of demand for the good X_r.[1]

4. The Linear Preference Scale

The preference scale of the individual is defined as *linear* in the special case where the marginal rates of substitution are ratios of linear expressions in the variable purchases:

$$\begin{aligned} R_1 : R_2 : R_3 : \ldots : 1 = {} & a_1 + a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n : \\ & a_2 + a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n : \\ & \ldots\ldots\ldots \\ & a_n + a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n . \end{aligned}$$

Suppose that the function ϕ, which is one form of the utility function index (1), can be written in quadratic form, i.e.

$$\begin{aligned} \phi(x_1, x_2, \ldots x_n) = {} & a_0 + 2a_1x_1 + 2a_2x_2 + \ldots + 2a_nx_n \\ & + a_{11}x_1^2 + a_{22}x_2^2 + \ldots + a_{nn}x_n^2 \\ & + 2a_{12}x_1x_2 + 2a_{13}x_1x_3 + \ldots \end{aligned}$$

Then

$$\begin{aligned} \tfrac{1}{2}\frac{\partial\phi}{\partial x_1} &= a_1 + a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n \\ \tfrac{1}{2}\frac{\partial\phi}{\partial x_2} &= a_2 + a_{12}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n \\ & \ldots\ldots\ldots \\ \tfrac{1}{2}\frac{\partial\phi}{\partial x_n} &= a_n + a_{1n}x_1 + a_{2n}x_2 + \ldots + a_{nn}x_n \end{aligned}$$

[1] See Hicks and Allen, *op. cit.*, p. 64.

The marginal rates of substitution are the ratios of these expressions. This case of a quadratic utility function thus leads to a linear preference scale. But, it should be noticed, the linear preference scale so obtained is symmetrical in form;

$$a_{rs} = a_{sr} \qquad (r \text{ and } s = 1, 2, \ldots n)$$

This is not the case for the linear preference scale in general. All that can be stated, therefore, is that, when the marginal rates of substitution are linear and symmetrical, the linearity may be due to the existence of a quadratic utility function. But the symmetrical form of the linear preference scale is unnecessarily restricted. In the general case of the linear preference scale we propose to consider, there is nothing to be gained by assuming a quadratic or any other particular utility function. The form taken by the utility function is of no importance; all we require is that the marginal rates of substitution should be of the special linear form.

The equilibrium conditions (2), in the case of the linear preference scale, reduce to the form:

$$e_1 + e_2 + \ldots + e_n = e$$

and
$$\begin{aligned} b_1 + b_{11}e_1 + b_{12}e_2 + \ldots + b_{1n}e_n \\ &= b_2 + b_{21}e_1 + b_{22}e_2 + \ldots + b_{2n}e_n \\ &= \quad \ldots\ldots\ldots \\ &= b_n + b_{n1}e_1 + b_{n2}e_2 + \ldots + b_{nn}e_n \end{aligned}$$

where
$$b_r = \frac{a_r}{p_r} \quad \text{and} \quad b_{rs} = \frac{a_{rs}}{p_r p_s}$$
$$(r \text{ and } s = 1, 2, \ldots n)$$

This set of n equations can be solved to give e_r as a function of e involving the co-efficients denoted by the b's. We write the equations

$$\begin{aligned} e_1 + \quad e_2 + \ldots + \quad e_n \quad - e &= 0 \\ b_{11}e_1 + b_{12}e_2 + \ldots + b_{1n}e_n - \lambda + b_1 &= 0 \\ b_{21}e_1 + b_{22}e_2 + \ldots + b_{2n}e_n - \lambda + b_2 &= 0 \\ \ldots\ldots\ldots \\ b_{n1}e_1 + b_{n2}e_2 + \ldots + b_{nn}e_n - \lambda + b_n &= 0 \end{aligned}$$

where λ denotes the common value of the n expressions in the b's previously written. Let

$$B = \begin{vmatrix} 1 & 1 & \dots & 1 & 0 \\ b_{11} & b_{12} & \dots & b_{1n} & 1 \\ b_{21} & b_{22} & \dots & b_{2n} & 1 \\ & & \dots\dots & & \\ b_{n1} & b_{n2} & \dots & b_{nn} & 1 \end{vmatrix}$$

Further, let B_r denote the co-factor of the rth element of the first row, and B_{rs} the co-factor of b_{rs} in the determinant B. All these determinants involve only the co-efficients b_{rs} (r and $s = 1, 2, \dots n$).

Making use of this determinant notation, the solution of the linear equations for the expenditure e_r can be written [1]

$$e_r = -\frac{1}{B}\left(-eB_r + \sum_{s=1}^{n} b_s B_{sr}\right) \qquad (r = 1, 2, \dots n)$$

i.e.

$$e_r = k_r e + c_r$$

where

$$k_r = \frac{B_r}{B} \quad \text{and} \quad c_r = -\sum_{s=1}^{n} b_s \frac{B_{sr}}{B}$$

In the case of the linear preference scale, therefore, the individual's expenditure function for any good reduces to the particularly simple linear form. The marginal expenditure on the good is now a constant k_r independent of the level of total expenditure. The average expenditure and the elasticity of expenditure, on the other hand, are still dependent on the total expenditure. They can, however, be given quite simple expressions. In fact,

$$w_r = k_r + \frac{c_r}{e}$$

and

$$\eta_r = \frac{1}{1 + \dfrac{c_r}{k_r e}}$$

[1] We also have $\lambda = \frac{1}{B}(-eB_{n+1} + \sum_{s=1}^{n} b_s B_{sn+1})$ where B_{n+1}, B_{1n+1}, B_{2n+1} ... are the co-factors of the elements of the last column in B. The solution can be checked by substituting these values of e_r and λ in the equations.

5. Independent Goods

In the general preference scale, the goods are defined as forming a completely independent set if the marginal rates of substitution, as expressed by the ratios

$$\frac{\partial u}{\partial x_1} : \frac{\partial u}{\partial x_2} : \ldots : \frac{\partial u}{\partial x_n}$$

are each a function of two variables only, the two variables of the partial derivatives of u involved. The marginal rate of substitution between any two goods depends on the purchases of these goods but not on the purchases of other goods.

Hence, the ratio of $\frac{\partial u}{\partial x_1}$ to $\frac{\partial u}{\partial x_2}$ is a function of x_1 and x_2 only, and the ratio of $\frac{\partial u}{\partial x_1}$ to $\frac{\partial u}{\partial x_3}$ is a function of x_1 and x_3 only. Further, the ratio of $\frac{\partial u}{\partial x_2}$ to $\frac{\partial u}{\partial x_3}$ is a function of x_2 and x_3 only, i.e. the variable x_1 disappears when the first two ratios are divided. It is, therefore, possible to write the marginal rates of substitution

$$\frac{\partial u}{\partial x_1}\bigg/\frac{\partial u}{\partial x_2} = \frac{\phi_1(x_1)}{\phi_2(x_2)}$$

and

$$\frac{\partial u}{\partial x_1}\bigg/\frac{\partial u}{\partial x_3} = \frac{\phi_1(x_1)}{\phi_3(x_3)}$$

where ϕ_1, ϕ_2 and ϕ_3 are single-variable functions. Extending the argument to all ratios, we have the following expression for the marginal rates of substitution in the independent goods case:

$$R_1 : R_2 : \ldots : R_{n-1} : 1 = \frac{\partial u}{\partial x_1} : \frac{\partial u}{\partial x_2} : \ldots : \frac{\partial u}{\partial x_n}$$
$$= \phi_1(x_1) : \phi_2(x_2) : \ldots : \phi_n(x_n)$$

One form of the utility function index for independent goods is obtained at once:

$$u = \overline{\phi}_1(x_1) + \overline{\phi}_2(x_2) + \ldots + \overline{\phi}_n(x_n)$$

where $\overline{\phi}_r(x_r) = \int \phi_r(x_r)dx \qquad (r = 1, 2, \ldots n)$

In one form of the index, therefore, the variables $x_1, x_2, \ldots x_n$ appear independently and additively. The goods $X_1, X_2, \ldots X_n$ can be regarded, in this limited sense, as contributing independently to utility in consumption.

It remains to consider the case of the preference scale which is both linear and independent. In this case,

$$R_1 : R_2 : \ldots : R_{n-1} : 1$$
$$= a_1 + a_{11}x_1 : a_2 + a_{22}x_2 : \ldots : a_n + a_{nn}x_n$$

and the equations determining the equilibrium purchases are

$$e_1 + e_2 + \ldots + e_n = e$$
$$b_1 + b_{11}e_1 = b_2 + b_{22}e_2 = \ldots = b_n + b_{nn}e_n$$

Eliminating $e_2, e_3, \ldots e_n$, we have

$$e = e_1 + \frac{1}{b_{22}}(b_1 - b_2 + b_{11}e_1) + \frac{1}{b_{33}}(b_1 - b_3 + b_{11}e_1)$$
$$+ \ldots + \frac{1}{b_{nn}}(b_1 - b_n + b_{11}e_1)$$

i.e.
$$e = b_{11}e_1K + b_1K - L$$

where
$$K = \sum_{r=1}^{n} \frac{1}{b_{rr}} \quad \text{and} \quad L = \sum_{r=1}^{n} \frac{b_r}{b_{rr}}$$

So
$$e_1 = \frac{1}{b_{11}K}e + \frac{1}{b_{11}}\left(\frac{L}{K} - b_1\right)$$

A similar result holds for the expenditure on any good. Hence, in the present case, we have the linear expenditure relation

$$e_r = k_r e + c_r$$

previously obtained, but the co-efficients are now in the form

$$k_r = \frac{1}{b_{rr}K} \quad \text{and} \quad c_r = \frac{1}{b_{rr}}\left(\frac{L}{K} - b_r\right)$$

It follows that

$$b_{rr} = \frac{k_1}{k_r}b_{11}$$

and
$$b_r = b_1 + b_{11}k_1\left(\frac{c_1}{k_1} - \frac{c_r}{k_r}\right)$$

Hence, if we have the linear and independent case and if the values of the k's and the c's have been found from budget data for all goods, then the values of all the co-efficients denoted by the b's can be deduced in terms of the values b_1 and b_{11}. The co-efficients b_{rr} can thus be placed in order of numerical magnitude and so can the differences $(b_r - b_1)$. These results are of no value, however, unless we have reason to postulate that the goods form a completely independent set.

6. Variation of the Preference Scale

It is now assumed that the preference scale of each individual of a given group is of the special linear form, but that the scale varies in other respects from individual to individual. It can also be assumed, in addition, that the market prices at which the individuals make their purchases are different for the different individuals of the group. For each individual, therefore, there is a set of co-efficients of the kind denoted by the b's of the above analysis. The values of these co-efficients vary from individual to individual on account of the variation of the preference scales and also on account of the variation of the market prices appropriate to the individuals. It is required to investigate the effect of this variation on the observed expenditures of the individuals in the group.

As a first case, we can suppose that the variation is limited to the co-efficients denoted by $b_1, b_2, \ldots b_n$, the other co-efficients being the same for all individuals. Write the values of the co-efficients in the case of the tth individual

$$\bar{b}_r + {}_tb_r \qquad (r = 1, 2, \ldots n)$$

where ${}_tb_r$ is the difference between the actual value and the average value $\bar{b}_r$ for all individuals of the group. The value of the determinant B, as used above, is thus the same for all individuals, as are the values of the co-factors of B. The constant k_r of the linear expenditure relation for any good has the same value for all individuals. Only the constant c_r varies from individual to individual. We can thus write the linear expenditure for the tth individual in the form

$$e_r = k_r e + c_r + {}_tv_r$$

where $k_r = B_r/B$ is the same for all individuals, where

$$c_r = -\sum_{s=1}^{n} \bar{b}_s \frac{B_{sr}}{B}$$

is the average of the various c_r's for different individuals and where ${}_tv_r$ is the residual appropriate to the tth individual. It is clear that ${}_tv_r$, which is the difference between the actual expenditure of the tth individual and the expenditure given by the average expenditure relation ($e_r = k_r e + c_r$), can be expressed

$$_tv_r = -\sum_{s=1}^{n} {}_tb_s \frac{B_{sr}}{B}$$

This residual, being a linear expression in the variations ${}_tb_r$ for the

various goods, is distributed normally if each of the variations ${}_t b_r$ is distributed normally over the individuals of the group, or if the variations ${}_t b_r$ for different goods are independent and the number of goods large. The expenditure lines for any good are parallel for the different individuals and the vertical intercepts of the lines are distributed normally under the conditions stated.

The residuals ${}_t v_r$ are also normally distributed over all individuals *with the same expenditure e* under the same conditions, provided that there is, in addition, no correlation between the expenditures of the individuals and the variations ${}_t b_r$. The standard deviations of these distributions are identical for all levels of expenditure. The distribution of the residuals is, in fact, homoscedastic.

The same distribution of residuals is found *approximately* in the general case where all the b co-efficients vary from individual to individual, provided that the variations in the co-efficients b_{11}, b_{22}, b_{12}, . . . are all small compared with the variations in the coefficients b_1, b_2, . . . The values of the co-efficient k_r for any good are not now exactly the same for all individuals, but the variation is much less marked than that of the values of c_r. The expenditure lines for any good are now nearly parallel for all individuals. The vertical intercepts are distributed approximately according to the normal law, about the average position of the line over all individuals, if conditions similar to those stated above hold.

7. Price Elasticities of Demand and Elasticities of Substitution

The equilibrium conditions (2) determine the demand of an individual for any good as a function of the individual's total expenditure and of all the market prices. It has been assumed, so far, that the prices are fixed so that total expenditure is the only variable. The variation of the demand for X_r with respect to variation in total expenditure is then described by means of concepts such as η_r, the income elasticity of demand for the good.

The variation of individual demand as prices vary, total expenditure being fixed, is described by means of an entirely different set of concepts, the price elasticities of demand. If the price p_r of the good X_r varies, we define the elasticity of demand for X_r with respect to its own price as

$$\eta_{rr} = \frac{p_r}{x_r} \frac{\partial x_r}{\partial p_r}$$

and the elasticity of demand for any other good X_s with respect to the price of X_r as

$$\eta_{rs} = \frac{p_r}{x_s}\frac{\partial x_s}{\partial p_r} \qquad (r \neq s)$$

The first can be called a "direct" price elasticity and the others "cross" price elasticities of demand.[1]

Suppose that the price of X_r varies while all other prices are fixed. In the case of the linear preference scale, to which we are confining our attention, the equilibrium conditions (2) can be written

$$\begin{aligned} p_1x_1 + p_2x_2 + \ldots + p_nx_n &= e \\ a_1 + a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= \lambda p_1 \\ a_2 + a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= \lambda p_2 \\ &\ldots\ldots\ldots \\ a_n + a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n &= \lambda p_n \end{aligned}$$

where λ denotes the common equilibrium value of the ratios of the marginal rates of substitution to the corresponding prices. Differentiate these equations with respect to p_r and note that

$$\frac{\partial x_r}{\partial p_r} = \frac{x_r}{p_r}\eta_{rr} \quad \text{and} \quad \frac{\partial x_s}{\partial p_r} = \frac{x_s}{p_r}\eta_{rs}$$

Then

$$\begin{aligned} p_1\frac{x_1}{p_r}\eta_{r1} + p_2\frac{x_2}{p_r}\eta_{r2} + \ldots + p_n\frac{x_n}{p_r}\eta_{rn} &= -x_r \\ a_{11}\frac{x_1}{p_r}\eta_{r1} + a_{12}\frac{x_2}{p_r}\eta_{r2} + \ldots + a_{1n}\frac{x_n}{p_r}\eta_{rn} &= p_1\frac{\partial\lambda}{\partial p_r} \\ a_{21}\frac{x_1}{p_r}\eta_{r1} + a_{22}\frac{x_2}{p_r}\eta_{r2} + \ldots + a_{2n}\frac{x_n}{p_r}\eta_{rn} &= p_2\frac{\partial\lambda}{\partial p_r} \\ &\ldots\ldots\ldots \\ a_{r1}\frac{x_1}{p_r}\eta_{r1} + a_{r2}\frac{x_2}{p_r}\eta_{r2} + \ldots + a_{rn}\frac{x_n}{p_r}\eta_{rn} &= \lambda + p_r\frac{\partial\lambda}{\partial p_r} \\ &\ldots\ldots\ldots \\ a_{n1}\frac{x_1}{p_r}\eta_{r1} + a_{n2}\frac{x_2}{p_r}\eta_{r2} + \ldots + a_{nr}\frac{x_n}{p_r}\eta_{rn} &= p_n\frac{\partial\lambda}{\partial p_r} \end{aligned}$$

[1] If the demand for X_r falls as its price rises, the elasticity η_{rr} defined here is negative. The value adopted here is opposite in sign to that usually taken (after Marshall). It is convenient, however, to have a uniform notation for all elasticities.

Substitute

$$e_r = p_r x_r, \quad a_{rr} = p_r^2 b_{rr} \quad \text{and} \quad a_{rs} = p_r p_s b_{rs}$$

and clear of fractions. We then obtain the following set of linear equations in $\eta_{r1}, \eta_{r2}, \ldots \eta_{rn}$:

$$\begin{aligned}
e_1\eta_{r1} + \quad e_2\eta_{r2} + \ldots + \quad e_n\eta_{rn} \qquad\qquad &= -e_r \\
b_{11}e_1\eta_{r1} + b_{12}e_2\eta_{r2} + \ldots + b_{1n}e_n\eta_{rn} - p_r\frac{\partial\lambda}{\partial p_r} &= 0 \\
b_{21}e_1\eta_{r1} + b_{22}e_2\eta_{r2} + \ldots + b_{2n}e_n\eta_{rn} - p_r\frac{\partial\lambda}{\partial p_r} &= 0 \\
\cdots\cdots\cdots & \\
b_{r1}e_1\eta_{r1} + b_{r2}e_2\eta_{r2} + \ldots + b_{rn}e_n\eta_{rn} - p_r\frac{\partial\lambda}{\partial p_r} &= \lambda \\
\cdots\cdots\cdots & \\
b_{n1}e_1\eta_{r1} + b_{n2}e_2\eta_{r2} + \ldots + b_{nn}e_n\eta_{rn} - p_r\frac{\partial\lambda}{\partial p_r} &= 0
\end{aligned}$$

Solving these equations in determinant form, and making use of the determinant B and its co-factors as before, we obtain

$$e_r\eta_{rr} = -\frac{1}{B}(B_r e_r - \lambda B_{rr})$$

and
$$e_s\eta_{rs} = -\frac{1}{B}(B_s e_r - \lambda B_{rs})$$

So
$$\left.\begin{aligned}
-\eta_{rr} &= k_r + (1 - w_r)\sigma_r \\
-\eta_{rs} &= \frac{w_r}{w_s}k_s - w_r\sigma_{rs}
\end{aligned}\right\}$$

where
$$\sigma_r = -\frac{\lambda}{ew_r(1 - w_r)}\frac{B_{rr}}{B}$$

and
$$\sigma_{rs} = \frac{\lambda}{ew_r w_s}\frac{B_{rs}}{B}$$

The expression σ_r is defined as the elasticity of substitution between X_r and all other goods at the equilibrium position, and σ_{rs} as the partial elasticity of substitution of X_s for X_r at the equili-

brium position.[1] It can be shown, from the definition, that σ_r is not negative and that it indicates the degree to which substitution of other goods for X_r in consumption is possible. If $\sigma_r = 0$, no substitution is possible; if σ_r is large, other goods are readily substitutable for X_r. On the other hand, σ_{rs} can be positive, zero or negative; in the former case, X_s competes with X_r in consumption, and in the latter case, X_s complements X_r in consumption.

The numerical value of the direct price elasticity of demand for X_r $(-\eta_{rr})$ is the sum of two separate terms. The first term represents what can be called an "expenditure effect" and is measured by the constant k_r of the linear expenditure relation for the good X_r. If the price p_r falls, for example, there is a consequent increase in the real income of the individual and hence an increase (except in the exceptional case where k_r is negative) in the demand for X_r. The second term involves the elasticity of substitution between X_r and all other goods and represents a "substitution effect." If p_r falls, the good X_r is cheaper relatively to the other goods and the demand for X_r will increase on account of the substitution of X_r for other goods in consumption. The two "effects" together make up the whole change in the demand for X_r consequent upon its price change.

It follows that, if $\sigma_r = 0$ and no substitution between X_r and other goods is possible, the value of $(-\eta_{rr})$ reduces to the constant k_r which measures the expenditure effect. Since k_r is less than unity and usually quite small, the demand for a good with no

[1] Relations of the kind written above in the case of the special linear preference scale can be obtained also in the general case of a preference scale of any form. The general results are

$$-\eta_{rr} = w_r\eta_r + (1 - w_r)\sigma_r \quad \text{and} \quad -\eta_{rs} = w_r\eta_s - w_r\sigma_{rs}$$

where η_r and η_s are the income elasticities of demand for the goods X_r and X_s, and where the elasticities of substitution σ_r and σ_{rs} take more general expressions than those written in the special case considered here. For an account of the definition of the elasticities of substitution (in terms of the general preference scale) and a derivation of the general results, see Hicks and Allen, *op. cit.* One change of notation and terminology, in addition to the change in the signs of the price elasticities, is introduced here. The partial elasticity of substitution σ_{rs} is opposite in sign to what was called, in the article referred to, the elasticity of complementarity of X_s with X_r. The statements made in the text here concerning the meanings of the elasticities σ_r and σ_{rs} are derived from the article.

Dr. J. R. Hicks, since the publication of the joint article, has elaborated the concepts of elasticity of substitution and has extended their application in a number of directions. His results will be published in a forthcoming book on *Value and Capital*. The change of notation adopted here has been made in order to bring the present development into line with his own. We are much indebted to him for suggestions to this effect.

substitutes is inelastic with respect to changes in its own price. As substitution becomes more possible, larger values of $(-\eta_{rr})$ can be obtained and the demand for the good can become elastic. In any case, the elasticity of demand must be small unless there are considerable substitutional relations between the good concerned and other goods.

The constant c_r of the linear expenditure relation for the good X_r can be expressed in terms of the elasticities of substitution between X_r and other goods. We have

$$c_r = -\sum_{s=1}^{n} b_s \frac{B_{sr}}{B}$$

i.e.
$$c_r = \frac{w_r e}{\lambda}\{b_r(1 - w_r)\sigma_r - \Sigma b_s w_s \sigma_{sr}\}$$

where the summation of the second term is for all values of s from 1 to n except the value $s = r$.

Hence, in interpreting the magnitude of the constant c_r for any good, the values of the co-efficients $b_1, b_2, \ldots b_n$ and of the elasticities of substitution are factors of importance. It can be taken that all co-efficients $b_1, b_2, \ldots b_n$ are positive,[1] and, since σ_r, $(1 - w_r)$ and w_s are positive, c_r can only be negative if at least one of the σ_{sr}'s is positive. The good X_r can only be a luxury ($c_r < 0$) if some of the σ_{sr}'s are positive, i.e. if X_r is in competition with at least some of the other goods. On the other hand, if X_r is a necessary ($c_r > 0$), there is, in general, less competition between X_r and other goods and the more urgent is the good X_r the more pronounced must be the complementary relations between X_r and other goods. Since the c_r's for various goods add up to zero, the values of the σ_{sr}'s must be positive on balance in order to counteract the positive elasticities σ_r in the expressions for the c_r's. It is clear, therefore, that complementarity is the exception rather than the rule in the relationship of goods in consumption.

[1] It is assumed, in the normal form of the preference scale, that the marginal rates of substitution are positive. As the purchases x_1, x_2, x_n become smaller, in the case of the linear preference scale, the marginal rates of substitution tend to equal $b_1 : b_2 : \ldots : b_n$. The b's can thus be taken as positive.

Printed in Great Britain by Butler & Tanner Ltd., Frome and London